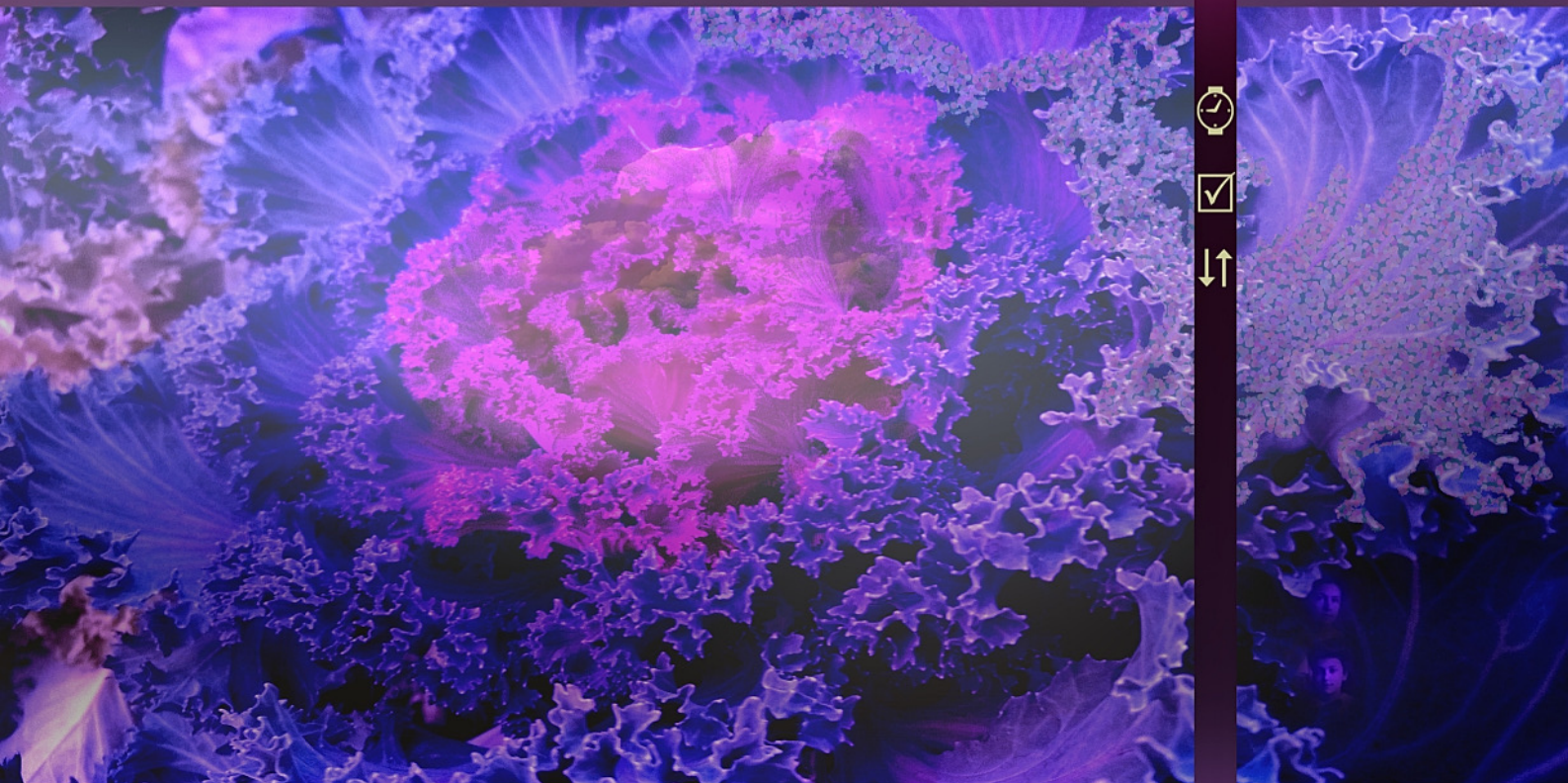


# RoCHI – International Conference on Human-Computer Interaction



**Proceedings of the 14<sup>th</sup> Conference  
on Human-Computer Interaction – RoCHI 2017**  
11—12 September 2017 • Craiova, Romania

## EDITORS

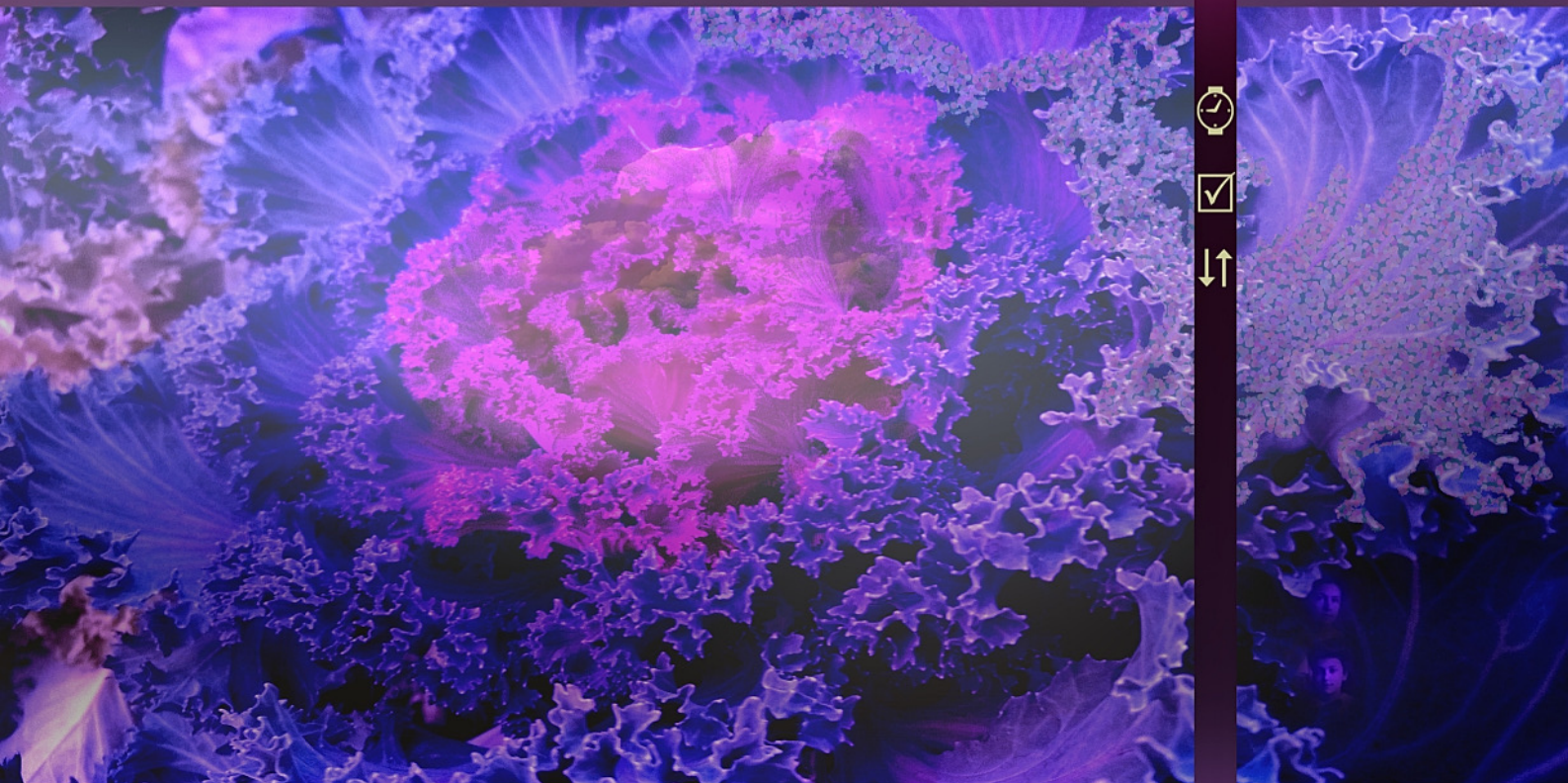
**Marian Cristian Mihaescu**  
University of Craiova

**Peter Forbrig**  
Universität Rostock





# RoCHI – International Conference on Human-Computer Interaction



**Proceedings of the 14<sup>th</sup> Conference  
on Human-Computer Interaction – RoCHI 2017**  
11—12 September 2017 • Craiova, Romania

## EDITORS

**Marian Cristian Mihaescu**  
University of Craiova

**Peter Forbrig**  
Universität Rostock



**ROCHI – INTERNATIONAL CONFERENCE ON  
HUMAN-COMPUTER INTERACTION**

Proceedings of the 14th International Conference on Human-  
Computer Interaction RoCHI 2017, 11-12 September, Craiova,  
Romania

*Edited by:*

**Marian Cristian Mihăescu**

*University of Craiova, Romania*

*and*

**Peter Forbrig**

*University of Rostock, Germany*

MATRIX ROM  
Bucharest, 2017

Publisher

**MATRIX ROM**

C.P. 16-162

062510 – Bucuresti, Romania

Tel.: 021 4113617, Fax: 021 4114280

E-mail: office@matrixrom.ro

www.matrixrom.ro

MatrixRom publishing house is a certified publisher by the National Council for Scientific Research in Higher Education (Consiliul Național al Cercetării Științifice din Invățământul Superior)

Cover graphic design: *Sabin-Corneliu Buraga*, A.I.Cuza University, Iasi  
Graphic identity: *Adrian Mironescu*, Idegrafo

© Copyright 2017

All rights of the current edition are reserved by MATRIX ROM.

No part of this work may be reproduced or transmitted in any form or by any means, electronic or mechanical, without permission in writing from the publisher.

RoCHI – International Conference on Human Computer Interaction (2017)

<http://rochi.utcluj.ro/proceedings/en/index.php>

**ISSN 2501-9422**

**ISSN-L 2501-9422**



## TABLE OF CONTENTS

<b>Organization</b>	VI
<b>Foreword</b>	XI
<b>Invited Talks</b>	
• Contextual App'ification ..... <i>Christian Stary</i>	1
<b>Education</b>	
• LearnerSkill – A Serious Hex Wars Game Centered on Learner Comprehension ..... <i>Alexandru Grigoroiu, Irina Toma, Mihai Dascalu, and Stefan Trausan-Matu</i>	11
• Exploring the Perceived Ease of Use by Professors in Tesys e-Learning Platform ..... <i>Paul Stefan Popescu, Cristian Mihaescu, Mihai Mocanu and Costel Ionascu</i>	15
• Learn-o-Bot: Educational Application Using Augmented Reality and Sphero SPRK+ ..... <i>Andreea Camelia Tripon and Teodor Stefanut</i>	21
• Overview of virtual worlds usage for student engagement ..... <i>Stefan-Lucian Bucur, Florica Moldoveanu and Alin Moldoveanu</i>	27
<b>User-Centric Approaches</b>	
• Agile Software Development Process with Task Models having Triggers, Preconditions and Context Constraints ..... <i>Peter Forbrig and Gregor Buchholz</i>	35
• Active Aging in Place, a Motivational and Comprehensive Solution for a Better Longterm Care ..... <i>Adriana Alexandru, Marilena Ianculescu and Maria Gheorghe-Moisii</i>	43
• A user-centric approach to test automation of web-based applications ..... <i>Ionut Petre, Dragos-Marian Smada, and Radu Boncea</i>	49
• Predicting User Activities using GPS Collections ..... <i>Adrian Iftene</i>	53

## **Usability Evaluation and Natural Language Processing**

- A case study of usability evaluation: the Center for Active Ageing website ..... 61  
*Marilena Ianculescu, Ovidiu Bica, Dragos Daniel Iordache, and Costin Pribeanu*
- Using expert evaluation to assess the usability of the National Library of Software Programs ..... 65  
*Lidia Bajenaru, Ion Alexandru Marinescu, Costin Pribeanu, and Dragos Daniel Iordache*
- Rhythm analysis in chats using Natural Language Processing ..... 69  
*Irina-Diana Niculescu and Stefan Trausan-Matu*
- Using Text Processing in a Multimedia Environment. UAIC Activities in MUCKE Project ..... 75  
*Adrian Iftene*

## **Social Networks and Social Media**

- Perceived usefulness of Facebook for university students: A gender analysis across two countries ..... 81  
*Costin Pribeanu, Alexandru Balog, Vincentas Lamanauskas, and Violeta Slekiene*
- The social media impact on local e-government institutions in Romania ..... 87  
*Gheorghe-Moisii Maria, Tirziu Eugenia and Iordache Dragos-Danie*
- Identifying Similarities between Tweets ..... 91  
*Adrian Iftene*
- Identification of Agreement and Disagreement in Chat Conversations ..... 95  
*Cetin Rasid and Stefan Trausan-Matu*

## **Virtual and Augmented Reality**

- Gesture-based Visual Analytics in Virtual Reality ..... 101  
*Mihai Pop and Adrian Sabou*
- Considerations for a Virtual Training System to Improve Job Interview Skills for Software Engineers ..... 107  
*Iulia-Cristina Stanica, Florica Moldoveanu, Maria-Iuliana Dascalu, Alin Moldoveanu, and Dan Alexandru Mitrea*
- Tower Defense with Augmented Reality ..... 113  
*Adrian Iftene*
- Gamification in the context of user Productivity and engagement ..... 119  
*Sfat Lorena Adina and Stefanut Teodo*



## Interaction and Specification

- Conversational Agents Embodying a Character Using Deep Neural Networks ..... 125  
*Mihai Ilie and Traian Rebedea*
- A pipe-based approach for service-oriented UIs ..... 131  
*Mathias Kühn and Peter Forbrig*
- Interactive Visualization of Spatial Multidimensional Data ..... 135  
*Bogdan Maxim and Dorian Gorgan*
- SpiMO - Sitting Posture Monitoring System ..... 143  
*Jullia Birsan, Diana Stavarache, Maria-Iuliana Dascalu and Alin Moldoveanu*

## Techniques

- Transformation of the system sequence diagram to an interface navigation diagram ..... 147  
*William Germain Dimbisoa, Thomas Mahatody and Josvah Paul Razafimandimby*
- Mapping Questions to Ontology Components for Question Answering over Linked Data ... 155  
*Marius Valeriu Stanciu, Eugen Vasilescu, Stefan Ruseti and Traian Rebedea*
- Real-time Guided Procedural Terrain Generation ..... 159  
*Vlad Grigore Sima and Adrian Sabou*
- Improving an NP-Chunker by Exploiting the Inner NP Dependency Structure ..... 163  
*Mihaela Colhon and Dan Cristea*

## **ORGANIZERS**

### *Conference Chair*

**Marian Cristian Mihăescu**, University of Craiova, Craiova, Romania

### *Program Committee Chair*

**Peter Forbrig**, University of Rostock, Germany

### *Associate Chairs*

**Dorian Gorgan**, Technical University Cluj-Napoca, Romania

**Adrian Iftene**, A.I.Cuza University, Iasi, Romania

**Costin Pribeanu**, ICI Bucharest, Romania

**Stefan Trausan-Matu**, Politehnica University, Romania

### *Organizing Committee Chair*

**Marius Brezovan**, University of Craiova, Craiova, Romania

**Mihai Mocanu**, University of Craiova, Craiova, Romania

### *RoCHI Conference Website Administrator*

**Teodor Stefanut**, Technical University Cluj-Napoca, Romania

## **ORGANIZING COMMITTEE**

**Elvira Popescu**, University of Craiova, Craiova, Romania

**Mihaela Colhon**, University of Craiova, Craiova, Romania

**Costel Ionașcu**, University of Craiova, Craiova, Romania

**Cosmin Poteraș**, University of Craiova, Craiova, Romania

**Ștefan Popescu**, University of Craiova, Craiova, Romania

**Andrei Mocanu**, University of Craiova, Craiova, Romania



## **PROGRAM COMMITTEE**

**Alexandru Balog**, *ICI Bucharest, Romania*  
**Gilbert Cockton**, *Northumbria University, United Kingdom*  
**Sabin-Corneliu Buraga**, *Alexandru Ioan Cuza University, Iași, Romania*  
**Grigoreta Sofia Cojocar**, *Babeș-Bolyai University, Cluj-Napoca, Romania*  
**Marian Dârdală**, *Academy of Economic Studies of Bucharest, Romania*  
**Mihai Dascălu**, *University Politehnica of Bucharest, Romania*  
**Philippe Dessus**, *Université Grenoble Alpes, France*  
**Bruno Dumas**, *Université de Namur, Belgium*  
**Paul Fogarassy-Neszly**, *BAUM Engineering, Romania*  
**Peter Forbrig**, *University of Rostock, Germany*  
**Vivian Genaro Motti**, *George Mason University, Fairfax, United States*  
**Juan Gonzalez Calleros**, *Benemérita Universidad Autónoma de Puebla, Mexico*  
**Dorian Gorgan**, *Technical University of Cluj-Napoca, Romania*  
**Adriana-Mihaela Guran**, *Babeș-Bolyai University, Cluj-Napoca, Romania*  
**Adrian Iftene**, *Alexandru Ioan Cuza University, Iași, Romania*  
**Dragoș Daniel Iordache**, *ICI Bucharest, Romania*  
**Cristophe Kolski**, *Universite de Valenciennes*  
**Alin Moldoveanu**, *University Politehnica of Bucharest, Romania*  
**Philippe Palanque**, *IRIT, Universite Paul Sabatier, France*  
**Oscar Pastor**, *Polytechnic University of Valencia, Spain*  
**Lyn Pemberton**, *University of Brighton, United Kingdom*  
**Dorin Mircea Popovici**, *University Ovidius of Constanța, Romania*  
**Costin Pribeanu**, *ICI Bucharest, Romania*  
**Traian Eugen Rebedea**, *University Politehnica of Bucharest, Romania*  
**Adriana-Elena Reveiu**, *Academy of Economic Studies of Bucharest, Romania*  
**Cristian Rusu**, *Pontificia Universidad Católica de Valparaíso, Chile*  
**Carmen Santoro**, *ISTI-CNR Pisa, Italy*  
**Corina Sas**, *Lancaster University*  
**Ahmed Seffah**, *University of Lappeenranta, Finland*  
**Marcin Sikorski**, *Gdansk University, Poland*  
**Teodor Ștefănuț**, *Technical University of Cluj-Napoca, Romania*  
**Ștefan Trăușan-Matu**, *University Politehnica of Bucharest, Romania*  
**Jean Vanderdonckt**, *Université catholique de Louvain, Belgium*

## *ADDITIONAL REVIEWERS*

Felix Albu, Valahia University Targoviste, Romania  
Adriana Alexandru, ICI Bucharest, Romania  
Victor Bacu, Technical University Cluj-Napoca, Romania  
Florina-Oana Balan, Politehnica University, Romania  
Elena Bautu, Ovidius University, Romania  
Crenguta Bogdan, Ovidius University, Romania  
Tiberiu Boroş, Research Institute for Artificial Intelligence, Romania  
Gregor Buchholz, University of Rostock, Germany  
Alexandru Butean, Lucian Blaga University of Sibiu, Romania  
Dumitru-Clementin Cercel, Politehnica University, Romania  
Grigoreta Sofia Cojocar, Babes Bolyai University, Cluj-Napoca, Romania  
Mihaela Colhon, , University of Craiova, Romania  
Cesar Alberto Collazos, University of Cauca, Columbia  
Liviu-Adrian Cotfas, Academy of Economic Studies in Bucharest, Romania  
Titus Furtuna, Academy of Economic Studies in Bucharest, Romania  
Vivian Genaro Motti, George Mason University, United States  
Jens Gerken, University of Applied Sciences Gelsenkirchen, Germany  
Florin Girbacia, University of Brasov, Romania  
Juan Gonzales-Calleros, Benemerita Autonomous University of Puebla, Mexico  
Gabriela Grosseck, Vest University, Timisoara, Romania  
Mathias Kühn, University of Rostock, Germany  
Vincentas Lamanauskas, Siauliai University, Lithuania  
Alboaie Lenuta, *Alexandru Ioan Cuza University, Iaşi, Romania*  
Valentina Marinescu, University of Bucharest, Romania  
Cornelia Melenti, Technical University Cluj-Napoca, Romania  
Verginica Mititelu, Research Institute for Artificial Intelligence, Romania  
Delia Mitrea, Technical University Cluj-Napoca, Romania  
Florica Moldoveanu, Politehnica University of Bucharest, Romania  
Gabriel Neagu, ICI Bucharest, Romania  
Stefan Negru, *Alexandru Ioan Cuza University, Iaşi, Romania*  
Mihaela Ordean, Technical University Cluj-Napoca, Romania  
Marian Padure, Babes Bolyai University, Cluj-Napoca, Romania  
Jorge Perez Medina, Catholic University of Lovain, Belgium  
Elvira Popescu, University of Craiova, Romania  
Dan Rotar, University of Bacau, Romania  
Adrian Sabou, Technical University Cluj-Napoca, Romania  
Jean-Sébastien Sottet, Luxembourg Institute of Science and Technology, Luxemburg  
Emil Stanescu, ICI Bucharest, Romania



**Conference organized with the support of:**



University of Craiova, Romania

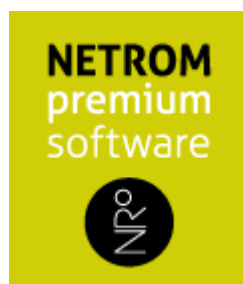


University of Rostock



RoCHI – ACM SIGCHI România

**This volume was published with the support of the:**



S.C. NETROM SOFTWARE S.R.L.



**CAPHYON**

S.C. CAPHYON S.R.L.

# Foreword

The International Conference on Human-Computer Interaction – RoCHI has reached its 14<sup>th</sup> edition this year. Like all the previous editions, RoCHI 2017 provides an inter-disciplinary forum for the exchange of ideas, expertise and research results in the field of HCI.

With this proceedings RoCHI is publishing two invited papers and 28 regular papers organized in seven thematic sessions. These papers were selected from a total of 40 submissions. After a careful review process, each submission being assigned to a minimum of three and maximum of six reviewers in a peer-reviewed fashion (double blind), a total number of 28 short and long papers were accepted, which represents an acceptance rate of 70%.

The invited paper by Chris Stary Johannes Kepler University Linz in Austria is entitled *Contextual App'ification – In this contribution, user-centered design and the generation of human support systems based on apps as behavior encapsulations are addressed from a conceptual and engineering perspective.*

The second invited talk is entitled *Empowering users with client-side Web augmentation tools and techniques - Mechanisms supported by client-side tools and techniques for allowing users to adapt Web sites according to their individual needs are described.* It is presented by Marco Winckler from Paul Sabatier University Toulouse.

The seven sessions of the conferences are entitled: Education, User-centric approaches, Usability evaluation and natural language processing, Social networks and social media, Virtual and augmented reality, Interaction and specification, and Techniques.

The papers were reviewed by an international program committee and some external experts. We express our appreciation to the members of the scientific community for selecting the best papers to be presented at the RoCHI 2017 conference. The reviewers provided very useful hints to the authors.

Finally, we acknowledge the efforts of all the persons involved in the organization of the RoCHI 2017 conference and thank them for their efforts!

Craiova and Rostock, July 20<sup>th</sup>

Marian Cristian Mihăescu

and

Peter Forbrig

# Contextual App'ification

Christian Stary

University of Linz

Business Informatics -  
Communications Engineering  
Christian.Stary@jku.at

## ABSTRACT

In this contribution, user-centered design and the generation of human support systems based on apps as behavior encapsulations are addressed from a conceptual and engineering perspective. Specific ingredients, in particular Complex Adaptive Systems (CAS) theory, a System-of-Systems (SoS) perspective, semantic interoperability of apps, and Subject-oriented Business Process Management (S-BPM) are detailed with respect to user-centered development. They allow the contextual alignment of application behavior from a role perspective, based on communication and interaction models that can be executed automatically. Users can set up the design space, validate specifications, and evaluate the context-sensitive alignment of apps. They primarily detail communication flows to meet their requirements and can experience the aligned apps interactively.

## Author Keywords

System(ic) Thinking, Complex Adaptive Systems; System-of-Systems; Subject orientation; User-centered Design; App'ification; Choreography.

## ACM Classification Keywords

D.2.10 Design: Methodologies; H.4.m. Information Systems Applications: Miscellaneous; H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

## General Terms

Human Factors; Design.

## INTRODUCTION

The design and development of interactive systems has changed significantly in the last few years, due to technological developments and resulting usage of applications (cf. [1][2][3][4]). Development has been driven by both, smart mobile devices, and dynamically adaptable application architectures. Using apps on mobile devices as intermediaries for accessing (cloud) services as standard interaction scenario leads to an app'ification of these services with a focus on user-specific needs. The term 'app'ification' or 'app-fication', respectively, has been used in several contexts, such as

- Personalization of (interactive) services [5]

- Security and privacy issues concerning interactive platforms (cf.[6])
- Speeding up of technology development and novel segmentation of system landscapes [7]
- Facilitating of third party developments for IoT (Internet of Things) platforms [8]
- Sustainability in a digital world and development of humanity when replacing human conversations, interactions and decisions through apps [9]

Gidey et al. [10] have been looking for architecturally significant requirements and architectural design decisions addressing several of the above mentioned issues. In their systematic elicitation study based on grounded theory they could identify 'appification' as '*the development and provision of smaller applications in a single feature set per application to work on devices with varying sets of platforms and form factors*' ([10], p.145). Following this understanding, a user-centered app'ification design approach requires dynamic composition and alignment facilities to adjust small applications to situation-sensitive needs and technological capabilities available in a specific situation.

As the number of apps a user utilizes increases the importance of contextual design involving various apps increases [8]. Consider a train information app. It can be of interest in the context of both, planning a trip, and being delayed on an ongoing trip. The latter requires real-time travel assistance in contrast to the first case which is based on planned schedules. From a user perspective, any difference in context should lead to different technical system behavior. Situation- or task-sensitive support requires dynamic design variants, albeit the need for consistent user experience of particular apps across multiple settings. Being able to adapt services in a context-dependent manner results in end-user computing, providing the capability to arrange apps according to an individual flow of control and dynamically changing requirements. Unlike the device-centric service computing the focus hereby is on a context-centric service experience caused by an app'ified design space. In such a space users should be able to design their digital services in a seamless manner.

Several challenges of contextual app'ification can be identified, when introducing a user-centered design approach, among them dynamic systems' thinking and

generating diagrammatic representations that can be executed under user control. Dynamic systems' thinking is based on understanding interactive systems as Complex Adaptive Systems which is introduced in the next section. It allows considering app'ification as System-of-Systems design process. Its structure is revealed in this section. System-of-Systems thinking gives space to create semantically interoperable designs which are populated with apps, and thus, triggers a respective development process (i.e. app'ification).

Semantic interoperability as essential part of app'ification is discussed in the subsequent section. This part of the paper also introduces process prototyping based on subject-oriented design representations, as it features task-specific user experience. Subject-oriented modeling allows mapping app-specific and cross-app behavior to executable system specifications, which in turn enables prototyping of app'ified support systems. The approach is exemplified for meeting specific needs from a user perspective. The final section concludes the contribution, summarizing the objectives and achievements, and providing inputs to further research.

#### TAKING A SYSTEM-OF-SYSTEMS PERSPECTIVE

In order to capture the dynamics of evolving socio-technical systems based on user needs, in this section relevant perspectives and concepts, namely Complex Adaptive Systems and System-of-Systems thinking, are provided.

#### Complex Adaptive Systems

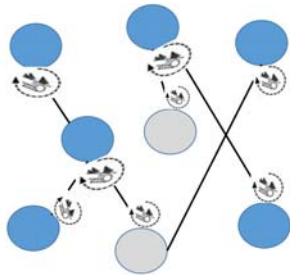
According to Chan [11] Complex Adaptive Systems (CAS) started in US to oppose the European "natural science" tradition in the area of cybernetics and systems. Although CAS theory shares the subject of general properties of complex systems across traditional disciplinary boundaries (like in cybernetics and systems) it relies on computer simulations as a research tool (as pointed out by Holland in 1992 initially [12]), and considers less integrated or "organized" systems, such as ecologies, in contrast to organisms, machines, or enterprises. Many artificial systems are characterized by apparently complex behaviors due to often nonlinear spatio-temporal interactions among a large number of component systems at different levels of organization, they have been termed Complex Adaptive Systems (CAS)

*CAS are dynamic systems able to adapt in and evolve with a changing environment. It is important to realize that there is no separation between a system and its environment in the idea that a system always adapts to a changing environment. Rather, the concept to be examined is that of a system closely linked with all other related systems making up an ecosystem. Within such a context, change needs to be seen in terms of co-evolution with all other related systems, rather than as adaptation to a separate and distinct environment ([11] p.2). CAS have several constituent properties (ibid. p.3ff):*

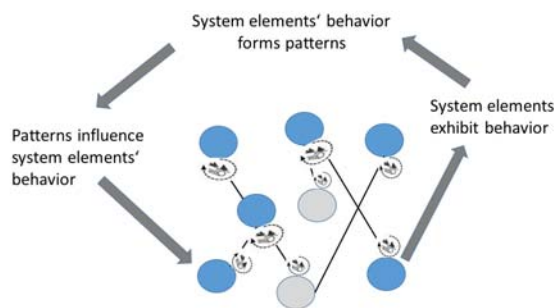
- *Distributed Control: There is no single centralized control mechanism that governs system behavior. Although the interrelationships between elements of the system produce coherence, the overall behavior usually cannot be explained merely as the sum of individual parts.*
- *Connectivity: A system does not only consist of relations between its elements, but also of relations with its environment. Consequently, a decision or action by one part within a system influences all other related parts.*
- *Co-evolution With co-evolution, elements in a system can change based on their interactions with one another and with the environment. Additionally, patterns of behavior can change over time.*
- *Sensitive Dependence on Initial Conditions: CAS are sensitive due to their dependence on initial conditions. Changes in the input characteristics or rules are not correlated in a linear fashion with outcomes. Small changes can have a surprisingly profound impact on overall behavior, or vice-versa, a huge upset to the system may not affect it. ... This means the end of scientific certainty, which is a property of "simple" systems (e.g., the ones used for electric lights, motors and electronic devices). Consequently, socio-technical systems, are fundamentally unpredictable in their behavior. Long-term prediction and control are therefore believed to not be possible in complex systems.*
- *Emergent Order: Complexity in complex adaptive systems refers to the potential for emergent behavior in complex and unpredictable phenomena. Once systems are not in an equilibrium they tend to create different structures and new patterns of relationships. ... Complex adaptive systems function best when they combine order and chaos in an appropriate measure – this phenomenon has been termed Far from Equilibrium. CAS in their dynamics combine both order and chaos, and thus, stability and instability, competition and cooperation, order and disorder – being termed State of Paradox.*

In Figure 1 a schema of complex socio-technical system as a group of different types of elements is shown, existing far from equilibrium, when forming interdependent, dynamic evolutionary networks that are sensitive dependent and fractionally organized (cf. [13]). Taking a CAS perspective requires system thinking in terms of networked, however modular elements acting in parallel (cf. [14]). In socio-technical settings, these elements can be individuals, technical systems or their features. Understood as CAS they form and use internal models to **anticipate** the future, basing current actions on expected outcomes. It is this attribute that distinguishes complex adaptive systems from other kinds of complex systems; it is also this attribute that

makes the emergent behavior of complex adaptive system intricate and difficult to understand ([12], p.24). Figure 2 visualizes the behavior-specific dynamics of CAS.



**Figure 1. A schema of a complex system.**



**Figure 2. A schema of a Complex Adaptive System.**

According to CAS theory, in CAS settings each element sends and receives signals in parallel, as the setting is constituted by each element's interactions with other elements. Actions are triggered upon other elements' signals. In this way, each element also adapts and thus, evolves through changes over time. Self-regulation and self-management have become crucial assets in dynamically changing socio-technical settings, such as organizations ([15][16]). Self-organization of concerned stakeholders as system elements is considered key handling requirements for change. However, for self-organization to happen, stakeholders need to have access to relevant information of a situation. Since the behavior of autonomous stakeholders cannot be predicted, a structured process is required to guide behavior management according to the understanding of stakeholders and their capabilities to change their situation individually [15] [17].

From the interaction of the individual system elements arises some kind of global property or pattern, something that could not have been predicted from understanding each particular element [11]. A typical emergent phenomenon is a Social Media momentum stemming from the interaction of the users when deciding upon a certain behavior, such as spontaneous meetings. Global properties result from the aggregate behavior of individual elements. Although it is

still an open question how to apply CAS to engineering systems with emergent behavior (cf. [14]), in case of socio-technical system design pre-programmed behavior is a challenging task, as humans may change behavioral structures in response to external or internal stimuli. As such, stakeholders in these systems (self-)organize evolvment and adapt to a changing environment, usually generating more complexity in the process.

### System-of-Systems

System-of-Systems (SoS) is considered an effective way of handling CAS, in particular when developing complex artefacts in a structured way [18]. According to IEEE's Reliability Society, thereby a system is a *group of interacting elements (or subsystems) having an internal structure which links them into a unified whole. The boundary of a system is to be defined, as well as the nature of the internal structure linking its elements (physical, logical, etc.). Its essential properties are autonomy, coherence, permanence, and organization* ([19], p.1). A System-of-Systems (SoS) is a system that involves several systems that are operated independently but have to share the same space and somehow cooperate (*ibid*, p.2).

As such, they have several properties in common: operational and managerial independence, geographical distribution, emergent behavior, evolutionary development, and heterogeneity of constituent systems (*ibid.*). These properties affect setting the boundaries of SoS and the internal behavior of SoS, and thus, influences methodological SoS developments ([20], p. 206). SoS are distinct with respect to

- (1) *autonomy where constituent systems within SoS can operate and function independently and the capabilities of the SoS depends on this autonomy,*
- (2) *belonging (integration), which implies that the constituent systems and their parts have the option to integrate to enable SoS capabilities,*
- (3) *connectivity between components and their environment,*
- (4) *diversity (different perspectives and functions),*
- (5) *emergence (foreseen or unexpected) (ibid.)*

Several structures and categorization schemes have been used when considering complex systems as System-of Systems, ranging from closed coupling (systems within systems) to loosely coupling (assemblage of system). They constitute embodied systems cooperating in an interoperable way (cf. [21]), allowing for autonomous behavior of each system while contributing through collaboration with other systems, in order to achieve the objective of the networked systems (SoS) [22] – see also next section.

Referring to structural and dynamic complexity, *structural complexity derives from i) heterogeneity of components*



across different technological domains due to increased integration among systems and ii) scale and dimensionality of connectivity through a large number of components (nodes) highly interconnected by dependences and interdependences. Dynamic complexity manifests through the emergence of (unexpected) system behavior in response to changes in the environmental and operational conditions of its components ([19], p.1).

A typical SoS example are contextualized apps available on a smartphone. Each of them can be considered as a system. When adjusting them along a workflow, e.g., to raise alert and guide a patient to the doctor, in case certain thresholds with respect to medical conditions are reached for a specific user, several of these systems, such as blood pressure app, calendar app, and navigation app, need to be coordinated and aligned for personal healthcare, updating the task manager of the involved users. In this case, the smartphone serves as SoS carrier, supporting the patient-oriented redesign of the workflow, and thus, the SoS structure. The apps of the smartphone can still be used stand-alone, while the smartphone serves as a communication infrastructure and provider of networked healthcare-relevant subsystems. It is the latter property qualifying the smartphone as carrier of a SoS.

#### APP'IFIED SYSTEM DESIGN

In order to align apps to meet user or situation requirements they need to be interoperable from a semantic perspective. In the following semantic interoperability is discussed in the context of System-of-Systems and dynamic adaptation. As subject-oriented models allow for CAS-specific behavior representations of apps, a corresponding development process can be defined. It is given in the second subsection, demonstrating modeling and process-driven prototyping.

#### Semantically Interoperable Systems

We recognize app'ification from a System-of-Systems (SoS) perspective, which requires recognizing context of app use to scope the SoS and the process of adaptation. Stakeholders need to *understand the whole system beyond its elements, sub-systems, assemblies and components, and recognize how each element / sub-system / assembly / component functions as part of the entire system. They are multifaceted, able to consider issues from a wide range of perspectives and points of view and possess a generalist's perspective* ([23], p.276). They also need to *understand the interconnections and the mutual influences and interrelations among system elements. Systems thinking involves thinking about the system's interactions, interrelationships, and interdependencies of a technical, social, socio-technical or multi-level nature* (ibid.).

In doing so, developers lay ground for emergent properties of systems, effecting perspectives beyond engineering an isolated system. A systemic representation, such as a SoS

specification, enables designing user-specific systems. Thereby, several constellations may occur ([24], p. 28).

- **Integrated Systems (not interoperable):** *The systems functionally depend on each other. Once one of the systems fails, the overall system fails.*
- **Integrated Interoperable Systems:** *The systems share the same meta-model. While remaining independent systems, the syntax and semantics of the exchanged information are well defined and may be exchanged seamlessly.*
- **Unified Interoperable Systems:** *Here an abstract layer is introduced in the communication. Referring to elements on that layer requires abstracting from individual details. However, the individual systems are free to modify their own syntax or semantics as long as it is possible to relate information to the abstraction layer commonly agreed upon.*
- **Federated Interoperable Systems:** *This is the loosest interoperable approach. It allows, but at the same time also requires, negotiating in an ad-hoc manner with respect to exchanged information syntax and semantics.*
- **Compatible but not interoperable Systems:** *Compatible systems exist next to each other, but no meaningful information is exchanged between them. These systems do not interfere with the other system. However, it is not possible to provide larger functionality through a meaningful combination of the individual system's functions.*

Although with respect to semantic interoperability, the state of the art is to use ontologies (cf. [25]), in app'ification systemic relationships require a more dynamic protocol [26]. It needs also to go beyond Domain Specific Languages (DSL) (cf. [27]), to become more agile and proactive. App'ified SoSs need to be described and represented as being autonomous while being interconnected with each other in order to contribute to a goal at a higher level.

Interoperability is required on the one hand to recognize autonomous system nature and on the other hand, at the same time to recognize emergent behavior due to systems' connectivity to other systems [28]. In case of autonomous systems kept isolated, neither connectivity nor diversification of behavior is enabled. As such, these system cannot be aligned explicitly according to a common goal, as required for designing semantically interoperable system [18]. The involved systems need to have the ability to cooperate with each other, in particular agreeing upon a common way of interaction in order to collaborate and share information (cf. [29]).

For instance, in case of interoperable home health care, task manager app entries can represent both, links to other apps

including transfer of data, e.g., blood pressure to be sent, and individual action items, e.g., reminders to perform a certain activity such as a yoga session. Interoperability is essential not only to overcome system isolation, but also to enable systems to be diverse and emergent in behavior. Hence, both, interaction patterns, behavior requires representation and are subject to structural change. Changes may occur on the level of individual systems, as well as on the level of interaction of systems, affecting the overall SoS behavior (cf. [30][31][21]). From the perspective of sharing data, a blog entry referring to blood pressure may be relevant not only for individual reflection on individual healthcare development, but also for adjusting medical treatment with physicians. As such, the blog entry could be part of a patient’s individual knowledge repository, and of a medical record of the patient.

Dynamic arrangements require knowing the structure *and* behavior of a system. Hence, in case blog entries should be shared with other systems, blogging does not only require a data format for entries and some structure to inherently arrange them for blogging, but also facilities for editing both to adapt to SoS, e.g., sorting along SoS categories like all tasks referring to blood pressure data.

The integrated representation of the structure and behavior of systems including their environmental conditions can either take into account tasks and related processes, e.g., Kolb et al. [32], Sary et al. [33], allowing for automated creation and adaptation of user interfaces for specific user roles. In case of mapping interaction sequences to workflow, e.g., as proposed by Franke et al. [34] deviations between modeled and actual behavior can be identified, triggering the re-design of technically systems in SoS. In any case, semantic interoperability requires context-sensitive understanding and representations (cf. [35][36]). Different contexts of use leads to different requirements induce a federated approach to SoS engineering. Context-relevant but implementation-independent representations allow capturing and resolving interoperability issues at design time [21].

### Subject-oriented Engineering

In this section we first provide the subject-oriented perspective on SoS development before discussing its application for contextual app’ification.

#### Subject-oriented modeling and execution support

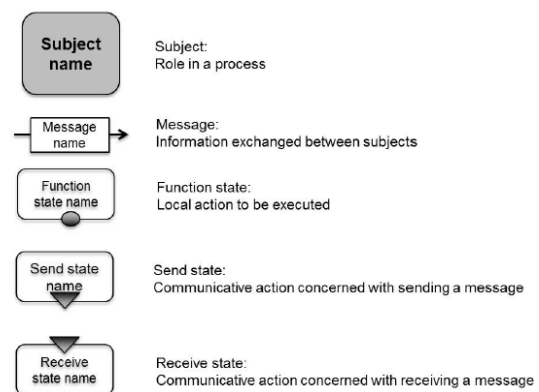
Subject-orientation is rooted in perceiving the world as parallel processes, each of which encapsulating a certain behavior and being able to trigger other processes and being triggered by other processes [37][38]. In subject-oriented SoS design systems are viewed as emerging from both the interaction between systems (represented as subjects) and their specific behaviors encapsulated within the individual systems [39]. According to this perspective, systems operate in parallel and can exchange messages, with the latter establishing a SoS.

The relations between systems of the SoS represent their context and scope the SoS. In that SoS, the systems operate autonomously and concurrently. When they are represented in a subject-oriented way systems are termed subjects. As behavior encapsulations they can be assigned to an entity that is capable of performing the encapsulated actions, from an organizational and/or technological perspective. This entity can be a human, an app, a machine (e.g., a robot), a device (e.g., a sensor), or a combination of these. Subject-oriented systems execute two different types of actions:

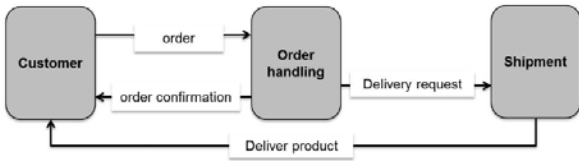
- *System-specific actions*: They correspond to self-contained activities and do not involve interacting with other systems, such as checking the blood pressure of a patient.
- *Communication with other systems*: These actions are concerned with exchanging messages between systems, i.e. sending and receiving messages, e.g., connecting a medical information system to the blood pressure measurement system of a patient, in order to receive medical feedback after collecting blood pressure data.

In subject-oriented representations, systems are represented as one of five core symbols listed in Figure 3. In Figure 4 subjects (i.e. systems) are Customer, Order Handling, and Shipment. Subject-oriented design requires two types of diagrams:

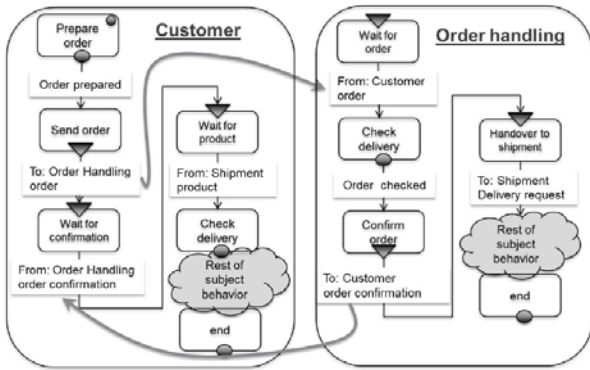
- *Subject Interaction Diagrams (SIDs)*. They provide a global view of a SoS, comprising its systems and the messages they exchange. The SID of a simple ordering setting is shown in Figure 4.
- *Subject Behavior Diagrams (SBDs)*. They provide an internal view on individual systems through sequences of states representing local actions (functions) and communicative actions (sending and receiving messages). State transitions are represented as arrows, with labels indicating the outcome of the preceding state (see Figure 5).



**Figure 3. Diagrammatic elements of subject-oriented specifications.**



**Figure 4. Handling orders – a SoS view via a Subject Interaction Diagram (SID).**



**Figure 5. System interaction for handling orders between Customer (SBD) and Order handling system (SBD).**

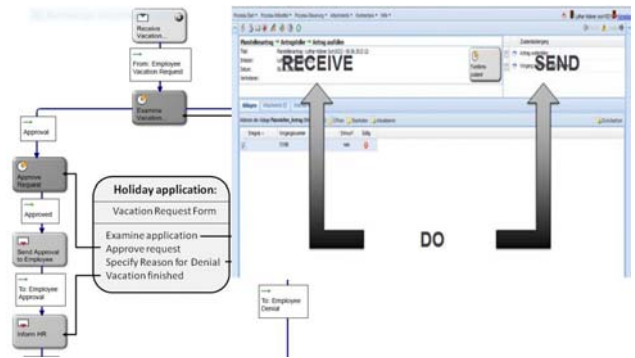
Given these notational capabilities SoS designs are characterized by

- a simple communication protocol (using SIDs) to relate systems
- standardized system interfaces (enabled by SBDs encapsulating system behavior)
- scalability in terms of complexity and scope, since each system can be decomposed according to functional or interaction patterns.

Subject-oriented representations allow representing each system's context in terms of communication relations to other systems. Hence, context-specific requirements, e.g., when creating a SoS for home health care, need to be mapped to relevant patterns of interactions between systems. The approach scales when SoS can be defined through dynamic and situation-sensitive system formations of systems in the sense of Complex Adaptive Systems (see section on CAS).

Since the communication relations between systems define a control flow for SoSs, validated subject-oriented specifications can be executed without further transformation (see Figure 6). In this way, users and stakeholders can be involved in modeling, refining, and implementing their SoS specifications.

Their modular structure is compatible with similar implementation approaches, such as service-oriented architectures and agent-oriented software systems [40].



**Figure 6. Interactive experience when executing subject-oriented system specifications.**

*Supporting context-sensitive SoS app'ification*

Subject-oriented SoS design seeks to assist user-centered system development by providing a methodology that presents behavior-relevant information encoded in apps. These characteristics enable users to be engaged more effectively via behavior representations. The SoS-based app'ification procedure has been designed according to the principles of subject-oriented system modeling and is structured as follows (cf. [41], Appendix):

- (1) An app is a system which is represented by its behavior.
  - (1.1) A specification of an app as system does not imply any actor or technology that could be used to execute the described behavior. It is implementation-independent.
  - (1.2) Apps can communicate with each other by exchanging messages (2).
- (2) A Message has a name and a payload.
  - (2.1) The name should express the meaning of a message informally.
  - (2.2) The payload corresponds to the data transported.
- (3) Apps have behavior representations. Internally,
  - (3.1) an app executes local activities
  - (3.2) an app sends messages to other apps
  - (3.3) an app expects messages from other apps
  - (3.4) an app performs all these activities in sequence – they are defined in an app's behavior specification.
- (4) Subject-oriented SoS specifications concern a specific behavior embedded in some context.
  - (4.1) Context is defined by the needs of users and properties of the situation at hand, including a set of apps (1,3).
  - (4.2) Context is expressed in terms of messages (2) for each app (1).
  - (4.3) Context is also provided by the technological infrastructure by which a system is technical part of, allowing to execute a specified app behavior (3).

Hence, contextual app'ification is based on the behavioral entities or abstract resources involved in a user-relevant setting. It is enabled and set up by systems termed subjects.

The behavior of a resulting SoS is determined by the subjects' exchanges of messages (i.e. system interactions).

(5) Modeling requires several activities, namely, the specification of

- (5.1) the scope of modeling
- (5.2) the apps to be processed for task completion
- (5.3) interactions the apps are part of
- (5.4) the messages they send or receive through each interaction
- (5.5) the behavior of each app encapsulating functions and interactions

(6) Subject Interaction Diagram (SID)

(6.1) (5.2) to (5.4) constitutes a Subject Interaction Diagram representing an app'ified SoS.

(6.2) A Subject Interaction Diagram is the most abstract diagrammatic level of describing how tasks are completed

(6.3) For each subject of a Subject Interaction Diagram a Subject Behavior Diagram (7) needs to be constructed for a complete and coherent app'ified SoS model.

(7) Subject Behavior Diagram (SBD)

(7.1) (5.5) for each subject constitutes a Subject Behavior Diagram.

(7.2) Once an app has been identified in (6), the behavior of each app can be defined in a subject-oriented way.

(7.3) An app's behavior is described by three states (send, receive, internal function) and transitions between these states. Hence, when specifying the behavior of each subject, a sequence of sending and receiving messages, and activities to be set for task accomplishment need to be represented.

(7.4) The subject-oriented description of an app defines the sequence of sending and receiving messages, or the processing of internal functions, respectively.

(7.5) A Subject Behavior Diagram is the most concrete diagrammatic level of adjusting apps, thus establishing a subject-oriented SoS.

(8) The states of a Subject Behavior Diagram represent operations.

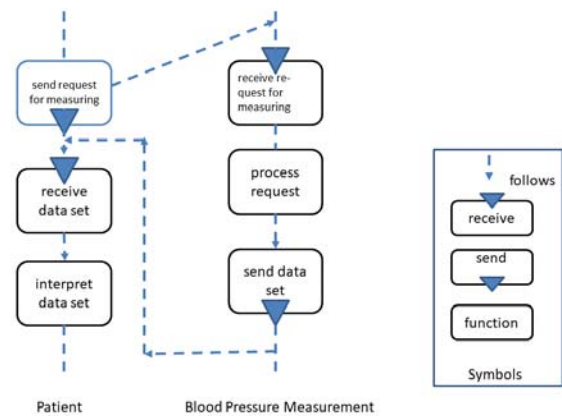
- (8.1) They are active elements of the subject description.
- (8.2) States are implemented by functions or services.
- (8.3) State transitions are necessary to exchange and manipulate data objects.

(9) Data objects are

- (9.1) data affected by operations of an app.
- (9.2) exchanged between apps by sending and receiving messages
- (9.3) processed through services.

In the following we exemplify a scheme how context-sensitive app'ification develops. We consider a home healthcare setting, as this kind of settings is under research increasingly from a socio-technical perspective (cf. [42]).

Consider the following case: An elder patient needs particular home healthcare support. Her blood pressure needs to be checked every other day. The patient needs to be reminded to trigger the measurement. Based on the results of the blood pressure check medical advice could be consulted. Three apps need to be aligned for effective monitoring and consultancy: (i) blood pressure measurement, (ii) medical expert consultancy, (iii) medication reminder & scheduling. In case the blood pressure reaches a certain threshold, medical experts need to be consulted, and eventually individual medication needs to be adjusted, triggering the reminder & scheduling system.



**Figure 7. SBD-home healthcare app'ification example involving Patient and Blood Pressure Management.**

In line with the subject-oriented SoS modeling each app is represented as individual system (subject), leading to a Subject Interaction Diagram involving 4 different systems and exchanging messages - in addition to the listed apps the patient also needs to be represented, in order to capture her behavior. As shown in the SBD interactions in Figure 7, the patient triggers the blood pressure measurement by sending a message to the corresponding subject (app), and receives a data set for interpretation.

Activating the medical advice could be modeled in a SBD through standard send-receive pattern with a medical consultancy system. The decision making procedure could also be added later, and the development requires CAS capabilities for app'ified SoSs. Putting the patient in control of triggering medical advice as dynamically occurring requirement means she triggers medical consultancy after checking her blood pressure, providing an up-to-date data set.

In order to mark such behavioral changes, and even adapt them further according to changing requirements, Fleischmann et al. [43,44] suggested the message guard concept for exception (i.e. non-routine behavior) and event handling. Events in subject-oriented representations occur

in a particular domain, such as home health care, and represent a novel need or a change in requirements. Events are interpreted as messages that can contain structured (data objects) or unstructured data (attachments).

The message guard reacting to the new behavior request is specified in Figure 8. Event producer, sending events in that case is the patient, requesting to decide by herself on a case-by-case basis. The event consumer is the message guard extending behavior capabilities. It receives events and processes them. In the example it may lead to involving medical advice, when the patient feels not comfortable with her interpretation of blood pressure date (see right branch). A function of a SBB allowing for behavior extension is marked with a triangle – see ‘interpret data set’ in Figure 8.

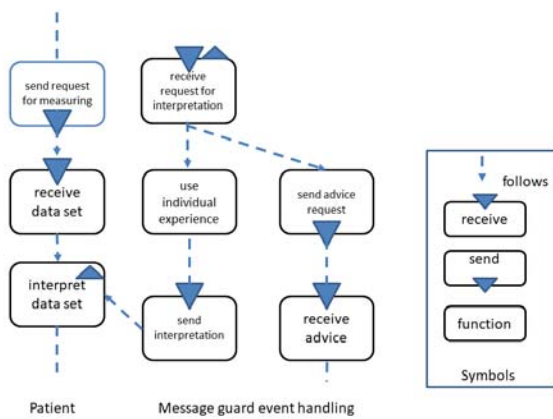


Figure 8. Capturing dynamic user control on medical advice.

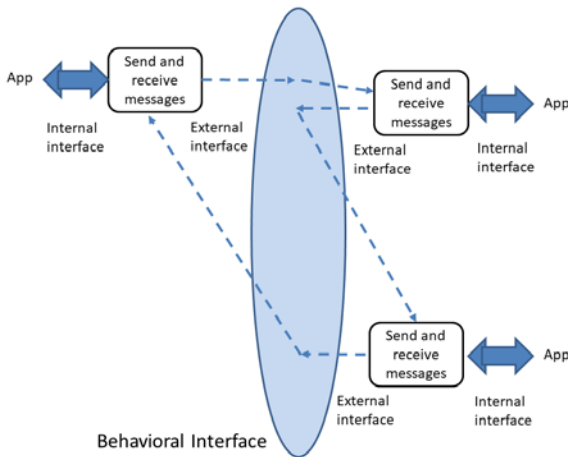


Figure 9. The Behavioral Interface enabling cross-app and cross-platform communication.

Utilizing subject-orientation, message exchange is the core feature for SoS app’ification including dynamic adaptation. For cross-app and cross-platform communication a message-based middleware could provide effective runtime support. Meyer et al. [45, 46] have developed a corresponding framework as shown in Figure 9. Each app

has an external and internal interface. The external one could be part of a routing component allowing to interact with other apps while the internal behavior is not visible to other apps. The Behavioral Interface keeps track, in which sequence messages are exchanged between apps, even across platforms. In this way also events can be exchanged, triggering behavior modifications, and thus, ad-hoc changes while maintaining semantic interoperability of the involved apps (cf. [47]).

### CONCLUSION

Digitalization of goods and processes enables informed design of services and systems by involved stakeholders, once they are provided with user-centered design instruments. Following a communication-driven perspective, system specifications can not only encapsulate behavior of apps, but also capture their context in terms of message exchanges. The underlying adaptive system-of-systems concept enables contextual app’ification of interactive systems, including the behavior of situation-or task-relevant actors, such as doctors, care takers, administration, and patients involved in home healthcare.

Subject-oriented representations contain functional and interactional aspects, and thus, contextual parameters to capture a situation relevant for system use. They can be used to generate interactive experience without touching the functionality of involved apps, rather extending their scope due to interactional alignment. At design and runtime semantic interoperability of involved systems can be ensured by context-sensitive behavioral interfaces.

Future research will, on the one hand focus on usability and acceptance of model-based design instruments for potential users, since, so far situation- or task-sensitive information has been captured. On the other hand, future work will focus on increasing the intelligence of app’ified systems, as the existing structure of aligned apps can be changed dynamically. Activities could then be proposed by system-of-systems, in particular referring to re-occurring behaviors or reference models. The more digital footprints become available, the more context can be considered for re-designing complex systems, such as Internet of Things or cyber-physical systems (cf. [48]).

### REFERENCES

1. Burroughs, B. YouTube kids: The app economy and mobile parenting. *Social Media+Society* 3, 2 (2017), DOI: <https://doi.org/10.1177/2056305117707189>.
2. Mansour, E., Sambra, A.V., Hawke, S., Zereba, M., Capadislis, S., Ghanem, A., Abounnaga, A. and Berners-Lee, T. A demonstration of the solid platform for social web applications. *Proc. of the 25<sup>th</sup> International Conference Companion on World Wide Web, International World Wide Web Conferences Steering Committee* (2016), 223-226.



3. Manresa-Yee, C. and Amengual, E. Tailoring ISO/IEC 12207 for Usability Engineering. *International Journal of Engineering Education* 32, 2 (2016), 886-893.
4. Magües, D.A., Castro, J.W. and Acuña, S.T. Requirements Engineering related Usability Techniques adopted in Agile Development Processes. *Proc. SEKE 2016*, Knowledge Systems Institute (2016), 537-542.
5. Shelley, C. Social agendas. In *Design and Society: Social Issues in Technological Design*. Springer International Publishing (2017), 105-124.
6. Fahl, S., Harbach, M., Perl, H., Koetter, M., and Smith, M. Rethinking SSL development in an appified world. *Proc. of the 2013 ACM SIGSAC conference on Computer & communications security*, ACM Press (2013), New York, 49-60.
7. Batran, A., Erben, A., Schulz, R., and Sperl, F. *Procurement 4.0: A survival guide in a digital, disruptive world*. Campus Verlag, Frankfurt (2017).
8. Jia, Y. J., Chen, Q.A., Wang, S., Rahmati, A., Fernandes, E., Morley, M. and Prakash, A. ContextIoT: Towards providing contextual integrity to appified IoT platforms. *Proc. of the 21<sup>st</sup> Network and Distributed System Security Symposium (NDSS'17)*, Internet Society (2017), 1-15.
9. Leonhard, G. and von Kospoth, C.A.G. Exponential technology versus linear humanity: Designing a sustainable future. In *Sustainability in a Digital World*, Springer International Publishing (2017), 77-83.
10. Gidey, H. K., Marmsoler, D. and Eckhardt, J. Grounded Architectures: Using Grounded Theory for the design of software architectures. *Proc. IEEE International Conference on Software Architecture Workshops (ICSAW)*, IEEE Computer Society (2017), IEEE Computer Society Press, New York, 141-148.
11. Chan, S. Complex adaptive systems. *ESD 83 Research Seminar in Engineering Systems*, 31 (2001).
12. Holland, J.H. Complex adaptive systems. *Daedalus* 121, 1 (1992), 17-30.
13. Fichter, Lynn S., Pyle, E.J. and Whitmeyer, S.J. Expanding Evolutionary Theory Beyond Darwinism with Elaborating, Self-Organizing, and Fractionating Complex Evolutionary Systems. *Journal of Geoscience Education* 58, 2 (2010), 58-64.
14. Holland, J. H. (2006). Studying complex adaptive systems. *Journal of Systems Science and Complexity*, 19(1), 1-8.
15. Allee, V. Value-creating networks: organizational issues and challenges. *The learning organization* 16, 6 (2009), 427-442.
16. Firestone, J.M. and McElroy, M.W. *Key issues in the new knowledge management*. Routledge, New York (2003).
17. Stary, C. Non-disruptive knowledge and business processing in knowledge life cycles—aligning value network analysis to process management. *Journal of Knowledge Management* 18, 4 (2014), 651-686.
18. Jamshidi, M. (Ed.) *System of systems engineering: innovations for the twenty-first century* (Vol. 58). John Wiley & Sons, New York (2011).
19. IEEE-Reliability Society. Technical Committee on 'Systems of Systems'. *Systems-of-Systems*, White Paper, 5p., IEEE, IEEE Society Press (2014).
20. Jaradat, R. M., Keating, C. B. and Bradley, J. M. A. histogram analysis for system of systems. *International Journal of System of Systems Engineering* 5, 3 (2014), 193-227.
21. Stary, C. and Wachholder, D. System-of-systems support—A bigraph approach to interoperability and emergent behavior. *Data & Knowledge Engineering* 105 (2016), 155-172.
22. Maier, M. W. Research challenges for systems-of-systems. *Proc. International Conference on Systems, Man and Cybernetics, Vol. 4*, IEEE, IEEE Society Press (2005), 3149-3154.
23. Frank, M. (2012). Engineering Systems Thinking: Cognitive Competencies of Successful Systems Engineers. *Procedia Computer Science* 8 (2012), 273-278.
24. Weichhart, G. and Stary C. Traceable pedagogical design rationales for personalized learning technologies: an interoperable system-to-system approach. *International Journal of People-Oriented Programming* 3, 2 (2014), 25-55.
25. Weichhart, G., Stary, C. and Vernadat, F.B. Enterprise Modeling for the interoperable and knowledge-based enterprise. *International Journal on Production Research* 55 (2017), in press.
26. Zacharewicz, G., Diallo, S., Ducq, Y., Agostinho, C., Jardim-Goncalves, R., Bazoun, H., Wang, Z. and Doumeings, G. Model-based approaches for interoperability of next generation enterprise information systems: state of the art and future challenges. *Information Systems and e-Business Management* (2016).
27. Weichhart, G. and C. Stary. A Domain Specific Language for Organisational Interoperability. *Proc. On The Move Conferences and Workshops*, Springer (2015), 117-126.
28. Bezerianos, A. and McEwan, G. Presence Disparity in Mixed Presence Collaboration. *Ext. Abstracts CHI 2008*, ACM Press (2008), 35–37.
29. Curry, E. System of systems information interoperability using a linked dataspace. *Proc. 7<sup>th</sup> International Conference on System of Systems*

- Engineering*, IEEE, IEEE Society Press (2012), 101-106.
30. Boardman, J. and Sauser, B. System of systems — the meaning of. *Proc. International Conference on System of Systems Engineering*, IEEE, IEEE Society Press (2006), 118-123.
  31. Baldwin, W.B. and Sauser B. Modeling the characteristics of system of systems, *Proc. International Conference on System of Systems Engineering*, IEEE, IEEE Society Press (2009), 1–6.
  32. Kolb, J., Hübner, P. and Reichert, M. Automatically Generating and Updating User Interface Components in Process-Aware Information Systems. *Proc. On the Move to Meaningful Internet Systems*, Springer 7565 Lecture Notes in Computer Science (2012), 444–454.
  33. Sary, C. and Wachholder, D. Context Control of Interoperable Interactive Systems. *Proc. 17th Conference on Business Informatics* vol. 1, IEEE, IEEE Society Press (2015), 232-241.
  34. Franke, J., Charoy, F. and El Khoury, P. Framework for coordination of activities in dynamic situations. *Enterp. Inf. Syst.* 7, 1 (2013), 33–60.
  35. Panetto, H. and Cecil, J. Information systems for enterprise integration, interoperability and networking: theory and applications. *Enterp. Inf. Syst.* 7, 1 (2013), 1–6.
  36. Vernadat, F.B. Technical, semantic and organizational issues of enterprise interoperability and networking. *Ann. Rev. Control* 34, 1 (2010), 139–144.
  37. Fleischmann, A., Schmidt, W., Sary, C., Obermeier, S. and Börger, E. *Subject-oriented business process management*. Springer Publishing Company, Incorporated (2012).
  38. Fleischmann, A., Schmidt, W. and Sary, C. Subject-oriented business process management. In *Handbook on Business Process Management 2*, Springer, Berlin (2015), 601-621.
  39. Sary, C. System-of-Systems Design Thinking on Behavior. *Systems* 5(1), 3 (2017).
  40. Fleischmann, A., Kannengiesser, U., Schmidt, W. and Sary, C. Subject-oriented modeling and execution of multi-agent business processes. *Proc. of the 2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technologies Volume 02*. IEEE, Computer Society (2013), 138-145.
  41. Fleischmann, A., Schmidt, W. and Sary, C. *S-BPM in the wild: Practical value creation*. Springer Publishing Company, Incorporated (2015).
  42. Ackerman, M.S., Goggins, S.P., Herrmann, T., Prilla, M. and Sary, C. *Designing Healthcare that Works. A Socio-technical Approach*. Academic Press, Cambridge, MA (2017).
  43. Fleischmann, A., Schmidt, W., Sary, C. and Strecker, F. Nondeterministic events in business processes. *Proc. International Conference on Business Process Management*. Springer (2012), 364-377.
  44. Fleischmann, A., Schmidt, W. and Sary, C. Complex Event Processing in e-Services. *Proc. 9th IEEE International Conference on Developments in eSystems Engineering (DeSE)*. IEEE Computer Society Press (2016), 251-259.
  45. Meyer, N., Feiner, T., Radmayr, M., Blei, D., Fleischmann, A. Dynamic catenation and execution of cross organisational business processes-the jCPEX! approach. *Proc. International Conference on Subject-Oriented Business Process Management*. Springer (2010), 84-105.
  46. Meyer, N., Radmayr, M., Heininger, R., Rothschädl, T., and Fleischmann, A. Platform for Managing and Routing Cross-Organizational Business Processes on a Network Router. *Proc. S-BPM ONE-Learning by Doing-Doing by Learning*, ed. Schmidt, W., Springer (2011), 175-189.
  47. Rothschädl, T. Ad-hoc adaption of subject-oriented business processes at runtime to support organizational learning. *Proc. S-BPM ONE-Scientific Research*, ed. Sary, C., Springer, 104 Lecture Notes in Business Information Processing (2012), 22-32.
  48. Neubauer, M. and Sary, C. *S-BPM in the Production Industry*. Springer International Publishing (2017).

# ***LearnerSkill* – A Serious Hex Wars Game Centered on Learner Comprehension**

**Alexandru Marian Grigoroiu, Irina Toma, Mihai Dascalu, Stefan Trausan-Matu**

University “Politehnica” of Bucharest

313 Splaiul Independenței, București

grigoroiaalexandru@gmail.com, irina\_toma@rocketmail.com, mihai.dascalu@cs.pub.ro

stefan.trausan@cs.pub.ro

## **ABSTRACT**

The serious games sector has shown a significant increase in the last years while relating to the overall released video games. Available serious games that provide Natural Language Processing (NLP) facilities only cover very specific tasks and, in most cases, are outdated in terms of User Interface (UI) and user experience. Thus, despite their educational purpose, this makes them less attractive and less engaging. The game introduced in this paper is a novel serious game that is based on *ReaderBench*, a powerful NLP framework, and a ranking system derived from the True Skill algorithm. Our aim consists of creating an easily extensible and highly accessible game that can accommodate multiple comprehension-centered mini-games. The UI is represented as a simple map conquest game in which learners play different mini-games that assess their reading skills. A preliminary validation was conducted on a group of 5 users that were particularly impressed by the concept of the game and the ease of adding new mini-games in the learning cycle.

## **Author Keywords**

Serious games; Natural Language Processing; TrueSkill ranking algorithm; *ReaderBench* framework.

## **ACM Classification Keywords**

I.2.1 [Applications and Expert Systems]: Games.

I.2.7 [Natural Language Processing]: Discourse, Language parsing and understanding, Text analysis.

## **General Terms**

Natural Language; Web-based Interaction; Games; Language acquisition; Text analysis.

## **INTRODUCTION**

Serious games are a rather new approach in e-Learning [10], that keeps expanding with every new game and idea implemented. The domain of serious games is wide and expands from educational contexts, like games that exercise memory or previous knowledge, to the military field, where games are used for training [7]. The advantage of using these games is that they enable situations that are hard to create in real life.

From the educational field, this article concentrates especially on serious games that enhance students' ability to learn from a given text. These games enhance students' reading abilities and their capacity to understand texts. Also, games should be personalized according to each

student's intrinsic abilities and qualities, such as: interest in the field, prior knowledge, innate learning capabilities, but also on external factors such as learning materials and environment.

In the classroom, teachers experiment with different learning materials to help students better understand the given information and at the same time to keep them engaged. Despite their efforts, the additional materials can be hard to use and understand, or even outdated. Even though teachers can assess the textual complexity of the taught lessons and the reading comprehension of each student, this method is daunting and prompt to errors. A solution to this problem is integrating serious games in the learning environment. These learning games use automated tools, which evaluate users based on their reading capabilities, and provide tasks to users to complete and enhance their skills.

From an educational point of view, having a strong evaluation system is enough for creating a successful serious game. But from the user perspective, if the game is not appealing or nor motivating enough, it will not provide the expected results [4].

The game introduced in this paper, *LearnerSkill*, targets all the aforementioned items. It uses the *ReaderBench* [2] framework for evaluating students' answers to different text-based tasks and it provides a fun and competitive game environment in which users improve their skills.

## **LEARNERSKILL**

*LearnerSkill* is a serious game that helps both students and teachers. On one side, students develop their reading and comprehension skills of English language by playing engaging and competitive games. On the other side, it helps tutors evaluate learners' progress over time and decide the best suited materials for their classes relying on the reading comprehension of each learner.

Expert readers have different strategies when it comes to reading a novel text. They are able to oversee their level of understanding at any given point, and when facing a difficulty, they rely on specific procedures, called reading strategies. Reading strategies are the mental processes used by students, consciously or unconsciously [8], when trying to understand a text. [1]. While integrating advanced Natural Language Processing (NLP) techniques, *ReaderBench* is able to identify reading strategies

employed by learners in their verbalizations and assess their self-explanations or collaborative contributions within chat forums [3]. In this first implementation, *LearnerSkill* uses the *ReaderBench* module centered on identifying reading strategies employed by learners while self-explaining a given text.

### Game Flow

*LearnerSkill* supports two types of users, tutors and learners. Tutors provide reading materials and create games based on them, can distribute learners in groups, track their evolution, and also participate in certain parts of the game offering aid and support. Whilst learners can compete in games by challenging one another or playing games proposed by tutors, track their own evolution in their profile and see what players have a similar rank. Tutors can upload or register text-based materials in the game that, on the server-side, are called documents. Based on each document one or more challenges can be created. A challenge is a task that a learner has to complete and it can be either a self-explanation, a summarization or selection of the most important sentences (see Figure 1). All the latter educational tasks become different mini-games within *LearnerSkill*. In terms of game design, our aim was to ensure extensibility with regards to the ease of adding new mini-games in the hexagons, such as highlighting important sentences or providing contextual definitions.

After creating challenges, tutors can group them into fixtures. A fixture can be made of one or more unique challenges. On the client side, a fixture corresponds to a game, which is composed of one or more rounds (corresponding to a server-side challenge). This grouping flexibility offers tutors endless options and the possibility to define tasks in coordination with the class curriculum. It can also let them test different scenarios for their chosen reading materials in order to find the best suited combination for their students' level of comprehension. This flow is explained in Figure 1.

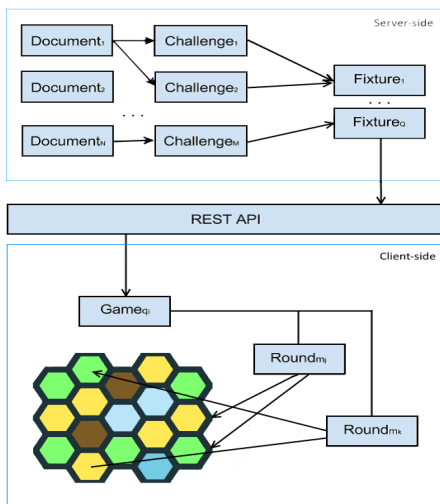


Figure 1. *LearnerSkill* game creation workflow.

This approach also offers the opportunity for tutors to track the answers submitted for a certain task. The capability to mix tasks that are based on different documents as part of the same fixture enables teachers to accurately evaluate how students act when faced with different combination of reading materials. In addition, the possibility of mixing tasks is not bound to materials covering a certain topic or subject. Depending on the materials teachers provide, any combination can be made, opening new perspectives on how students cope with reading on from different topics and of different difficulty levels, in various sequences.

### Ranking Algorithm

*LearnerSkill* was meant to create a competitive environment in which players not only compete against each other in a random fashion, but play with other students that match their skill. However, players are deliberately capable of choosing to compete against a certain user, if the tutor enables this facility.

TrueSkill [6] is the model chosen to rank users based on their skills as it is stable, feature rich, and thoroughly tested. The model is implemented in the JSkills library (available at <https://github.com/goochjs/JSkills>), which was afterwards adapted to fit our particular requirements.

The first difference is that *LearnerSkill* supports only head to head competitions and tasks, without any team based game modes. The second change addresses the limitation of the original TrueSkill implementation which accepts only an ordered hierarchy of players, i.e. it does not account for one's score, just the ordering within the final standings. Given the fact that *ReaderBench* offers a numerical value that evaluates the overall performance of a learner's submission, a simple strategy was devised to adapt *ReaderBench* evaluations to the TrueSkill implementation: in head to head mode, the evaluation for each player is compared and the best score wins.

Unfortunately, this simplistic approach makes it very difficult for draw situations to develop, as two players are highly unlikely to present the exact same level of comprehension for a given text. Having only win or loss outcomes may not necessarily represent an inconvenient; however, in time this may decrease the accuracy of the estimation of each learner's skill. Therefore, in order to produce match outcomes as draws, we decided to set a margin relative to the average score for each specific task. If the scores of both players fall within this margin, the game ends in a draw; otherwise one wins, whilst the other loses. Depending on the overall performance of players, these margins may be modified accordingly in order to offer the best game experience.

As *LearnerSkill* consists of games which are split into rounds, the learner who wins most rounds, also wins the game. It is important to note that the round or the game outcome does not necessarily matter in the economy of a learner's skill. The purpose of the game is not to rate how

well learners play it, but how well they understand texts, thus their comprehension skill. The outcome of the game is merely a motivational stimulus to continue playing the game while improve your personal skill.

The most important feature of the TrueSkill algorithm is represented by the ability to compute the draw margin for any given pair of players. This draw margin (or match quality) is computed using Formula 1:

$$\text{draw margin}(\beta^2, \mu_i, \mu_j, \sigma_i, \sigma_j) = \sqrt{\frac{2\beta^2}{2\beta^2 + \sigma_i^2 + \sigma_j^2}} e^{-\frac{(\mu_i + \mu_j)^2}{2(2\beta^2 + \sigma_i^2 + \sigma_j^2)}} \quad (1)$$

where  $\mu_i$  and  $\mu_j$  are the average value of skill,  $\sigma_i$  and  $\sigma_j$  are the standard deviations for two given players;  $\beta^2$  is the variance of performance and is given by the formula  $\beta^2 = (\sigma_0/2)^2$  [5]. For this algorithm, the original scale is kept given by a prior  $\mu_0 = 25$  and  $\sigma_0^2 = (25/3)^2$ .

Thus, our aim of ensuring an equitable game emerges in which leaners do not become stressed when facing only more experienced players or, in contrary, do not become demotivated when encountering only lower ranked learners. Moreover, our approach encourages active engagement while learners strive to achieve higher rankings that reflect more than basic cumulative scores.

### Graphical User Interface

The GUI is a simple web interface that makes use of the REST web services exposed by the *ReaderBench* framework. The current implementation targets the learner interface, while the tutor interface was limited to inserting the challenges directly in the database. After learners successfully sign in, they are presented with a hexagon grid, inspired from the Hex Wars games (available at <https://www.hexwar.com>), which represents the game map. The signed in learner is placed in the center tile of the grid. Other tiles of the map are occupied by other learners who participate in placed at a distance relative to their personal ranking. The closer the rank the other learners have, the closer they are placed to the center of the map.

In the current version of the UI, all learners start with a standard skill value at the game debut. In this situation, our system does not have any information on which learners could be best matched. An example of this situation can be seen in Figure 2. The signed-in user plays a game with another user (north-west position), both their skills are updated while the match quality for these two users changes. As the second user supposedly loses the game, his position changes and the corresponding modification is visible in Figure 2.b. As learners continue playing, their skills are updated accordingly, concurrently with their positions on the map.

From this map, the learner can choose an opponent to compete against. By hovering on the tile corresponding to another user, a tooltip element shows more information about that certain player, such as their score. By clicking on the other user's tile, the learner transitions to another

map, the *Conquest Map*, in which players are shown the rounds of a particular game, as set by the teacher. Once again, each round occupies its own tile. Completed rounds can be distinguished from available ones by tile coloring.

When selecting a round from a map, the learner is presented with the target text and a text box in which he can input his self-explanation. After users decide that the self-explanation is complete, they can submit it for evaluation to the *ReaderBench* service. Upon the completion of the automated analysis, *ReaderBench* responds with a score for each particular self-explanation and the round is marked as done.

The final result for the given round is not decided until both players have submitted their answers. In addition, learners are not required to be online at the same time in order to play. Mini-games can feature a limited time for the user to complete them, but this is optional and mostly dependent on tutor decisions. Skill updates occur only after the entire game has been decided and the average score from each round is computed. Based on this average, the outcome of a game can be either win, draw or loss. The result is fed to the ranking algorithm which computes the new ratings and updates the learners' skills. These changes are afterward reflected in the *Ranking Map* from Figure 2.b, as described in the earlier paragraphs.

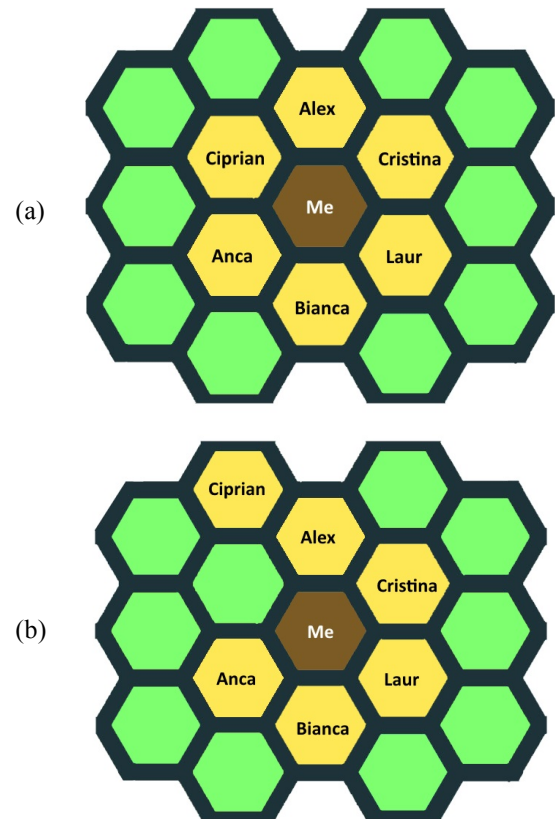


Figure 2. Ranking Map examples: a) at the start of the game; b) after skill update



## RESULTS

A preliminary validation was performed on a group of 5 students that were asked to play the game and provide feedback. They had to answer a 13 questions survey with ratings on a 5-point Likert scale (1—completely disagree; 5—completely agree). The questions were separated into three subject groups: a) general perspective, b) map design and usability and c) game interaction and flow. 90% of the users enjoyed playing the game and 80% considered it was clear that, the closer players are on the *Ranking Map*, the more competitive the game will be. Only 12% of the users did not understand that their skill is based on their performance in the task and not on the game outcome.

During the development phase, user feedback was also collected. Users suggested having a limited time for completing and submitting an answer, once the learner decides to answer a given task. Furthermore, answers should be checked against plagiarism, or, at least, users should not be able to paste text in the response box.

## CONCLUSIONS AND FURTHER DEVELOPMENT

*LearnerSkill* is an educational serious game that provides players with a competitive and entertaining environment for learning and assessing their comprehension skills. For teachers, it provides powerful tools to evaluate student performances in tight correlation with the provided learning materials. Moreover, teachers can further improve the quality of the learning materials, adapting them to better suit their students' level of comprehension.

The UI consists of a map conquest game in which players offer self-explanations. The submitted self-explanations are evaluated by *ReaderBench* and based on their score, the users may win, draw or lose, followed by a rating update.

In contrast to the previously developed serious games in our research group [9], *LearnerSkill* was specifically designed to accommodate extensibility in terms of rapid integration of new language mini-games, such as highlighting important phrases, extracting keywords, or providing contextual definitions similar to TOEFL testing. The potential mix of different tasks and learning activities creates the opportunity for greatly varying the educational scenarios. Moreover, the ranking algorithm can be applied on scores originating from any competitive game, thus ensuring a wide range of activities harmoniously integrated.

As further development, more mini-games will be included in the *LearnerSkill* engine, user feedback will be addressed and the overall user experience in the UI will be improved. In addition, we will include the possibility of a single player game mode in which students can complete tasks assigned by tutors, not necessarily in a competitive manner, but as homework.

## ACKNOWLEDGMENT

This work was partially funded by the 644187 EC H2020 *Realising an Applied Gaming Eco-system* (RAGE).

## REFERENCES

1. Dascalu, M., Dessus, P., Bianco, M., Trausan-Matu, S., and Nardy, A., 2014. Mining texts, learner productions and strategies with ReaderBench. In *Educational Data Mining: Applications and Trends*, A. Peña-Ayala Ed. Springer, Cham, Switzerland, 345–377.
2. Dascalu, M., Dessus, P., Trausan-Matu, S., Bianco, M., and Nardy, A., 2013. ReaderBench, an environment for analyzing text complexity and reading strategies. In *16th Int. Conf. on Artificial Intelligence in Education (AIED 2013)*, H.C. Lane, K. Yacef, J. Mostow and P. Pavlik Eds. Springer, Memphis, USA, 379–388.
3. Dascalu, M., Stavarache, L.L., Dessus, P., Trausan-Matu, S., McNamara, D.S., and Bianco, M., 2015. ReaderBench: An Integrated Cohesion-Centered Framework. In *10th European Conf. on Technology Enhanced Learning*, G. Conole, T. Klobucar, C. Rensing, J. Konert and É. Lavoué Eds. Springer, Toledo, Spain, 505–508.
4. Dempsey, K.B., Brunelle, J.F., Jackson, G.T., Boonthum, C., Levinstein, I.B., and McNAMARA, D.S., 2010. Miboard: Multiplayer interactive board game. *arXiv preprint arXiv:1009.2206*.
5. Herbrich, R., Minka, T., and Graepel, T., 2006. TrueSkill™: a Bayesian skill rating system. In *Proceedings of the 19th International Conference on Neural Information Processing Systems* MIT Press, 569-576.
6. Minka, T., Graepel, T.K., and Herbrich, R., 2013. Determining relative skills of players Google Patents.
7. Ritterfeld, U., Cody, M., and Vorderer, P., 2009. Serious games: Mechanisms and effects Routledge, New York, NY & London, UK.
8. Singhal, M., 2001. Reading proficiency, reading strategies, metacognitive awareness and L2 readers. *The Reading Matrix* 1, 1.
9. Toma, I., Andrioaie, I., Dascalu, M., and Trausan-Matu, S., 2016. Reading Space Secrets – A Serious Game Centered on Reading Strategies. *Revista Romana de Interactiune Om-Calculator* 9, 4, 269–284.
10. Westera, W., Nadolski, R.J., Hummel, H.G., and Wopereis, I.G., 2008. Serious games for higher education: a framework for reducing design complexity. *Journal of Computer Assisted Learning* 24, 5, 420–432.

# Exploring the Perceived Ease of Use by Professors in Tesys e-Learning Platform

Paul Ștefan Popescu<sup>1</sup>, Marian Cristian Mihăescu<sup>1</sup>, Mihai Mocanu<sup>1</sup>, Costel Ionașcu<sup>2</sup>

<sup>1</sup>Department of Computers and Information Technology

University of Craiova, Romania

{spopescu, mihaescu, mocanu}@software.ucv.ro

<sup>2</sup>Faculty of Economics and Business Administration

University of Craiova, Romania

icostelm@yahoo.com

## ABSTRACT

E-learning platforms are educational environments that have grown once with the information technology. They bring several features and advantages but there are also enough features that may be developed and others that may be explored. This paper presents an exploratory analysis regarding how the professor perceive the ease of use of the Tesys e-Learning platform. In order to perform the analysis, we use R and Wilcox statistics. We separated the data in two groups using the number of years and hours spent on the platform and tested them for group significant separation. The results show no difference between groups and so, we can conclude that all the subjects that completed the questionnaire are not influenced by other factors. There is also another implication of the results of this study because having the same results for both groups there is a consistency between all answers and the average grade a question got.

## Author Keywords

Interface evaluation; e-learning; HCI.

## ACM Classification Keywords

H.5.2. User Interfaces: *Graphical user interfaces (GUI); Interaction styles.*

## General Terms

Human Factors; Design; Measurement.

## INTRODUCTION

Once with the information technology growth alternatives to the classical learning environments appeared. E-learning platforms offer many benefits to learners because this environment offers access to learning resources without the need to be at a specific time in a specific place. Every professor has the possibility to make his own schedule but this comes with some drawbacks because the professors need to invest more time in learning resources in order to

offer better experience for learners. It is important for both students and professors to explore the perceived ease of use in order to be able to improve the e-learning experience. In our case we use Tesys [1] e-learning platform, an e-Learning platform developed at the University of Craiova which was designed for fulfilling specific needs but as new functionalities get implemented the usability in some cases decreases. One important aspect that needs to be mentioned is that the e-Learning platform was designed by computer professionals but used by people from different domains and the adaptability to the interface and functionalities tend to differ based on their context. In this paper, we aim to evaluate the professor's interface and functionalities in order to see how easy they adapt to the actual version of platform and also to gain a better intuition on what will be changed in the next release.

The interface of an e-Learning platform is very important for users that interact with it and as the platform receives more updates, some functionalities may become hard to be reached and also the new functionalities may not be intuitive enough.

The e-Learning platform used for this study is Tesys, a custom build platform that was entirely developed at the University of Craiova. There are four roles in which a user can be when he accesses Tesys e-Learning platform: student, professor, secretary or administrator. The biggest number of users is achieved by the role of student, then the role of professor has the next number of users and then secretaries which are two or three on the platform and the administrator which is only one. Based on this, we explored first the students' interface and now, this paper focuses on the professor's interface being the second as importance. As the number of professors and the level of usage grows, there are frequent questions regarding some functionalities but we think that there may be more functionalities that are not well optimized and we try to

discover the issues before the problems are reported. One other important issue is to explore the perceived ease of use and see what are the main factors that makes the differences between different groups.

E-learning platforms are evolving very fast and with new functionalities we need to explore how users perceives them. Professors that perform their action on e-Learning platforms are very important users because of their long-time activity in the platform and, they can offer a more relevant feedback. A better level of understanding regarding the professors' perceived ease of use will give an intuition regarding what can be changed or improved to offer a better learning experience.

In this paper, tested several factors that may influence the perceived ease of use in the e-Learning platforms. For the analysis procedure, we used Wilcoxon statistical method. As a expected result of this study there is knowledge that can be gained regarding the motives that makes Tesys e-Learning platform underutilized. The professors' attitude regarding the platform influence the interaction between user's and platform and also the attitude of students.

In this paper, we used the questionnaires method in order to find key issues that may be improved in the next release and to explore the ease of use perceived by the professors that teach their courses on the platform.

The questionnaires were administered to the professors that teach at the Faculty of Letters (FL) and the Faculty of Economics and Business Administration (FEAA). We used a total number of twenty-four questions divided in three categories: ease of use evaluation, control questions and group separation questions. In order to evaluate the perceived ease of use we used R as environment and Wilcoxon statistic method to explore the differences between groups.

#### **RELATED WORK**

Evaluation of e-Learning platforms [2] is a currently complex problem [3] due to many issues that need to be taken into consideration. Besides general issues like usability and accessibility, e-Learning platforms need to address the didactic effectiveness from student's and professor's perspective.

There are different approaches for evaluation of this kind of systems like eye movement approach [4]. This kind of approach tries to find areas where eye tracking may impact the interface evaluation but there is also a drawback because this kind of evaluation requires special equipment.

There are also papers that describe several other methods for usability evaluation in virtual environments like [5] in which they define and classify several types of evaluation

among with their issues. The authors state some distinctive characteristics of virtual environment characteristics like physical environments issues, evaluator issues, user issues, issues related to the type of evaluation and even other issues. After considering all those issues they define some current evaluation methods like Cognitive Walkthrough, Formative Evaluation, Heuristic or Guidelines-Based Expert Evaluation and others.

John P. Chin et. al. [6] conducted a research effort to develop the Questionnaire for User Interface Satisfaction (QUIS). For this study, the authors, had 150 pc users and he tested the reliability of the questionnaire using cronbach's alpha on two pairs of software categories: 1) software that was liked and disliked, and 2) a standard command line system (CLS).

Exploring how users perceives the ease of use and the technology acceptance is one important research issue that was addressed earlier than twenty-five years ago in [7].

This problem was further addressed even in a stronger relationship with e-Learning platform in other papers. One newer paper that explored the student's acceptance model in Jordan universities written by Amer Al- Adwan et. al. [8] explored the influence of perceived usefulness, perceived ease of use, attitude towards, and intention to use using Cronbach alpha [9]. This aggregation of metrics was called Technology Acceptance Model [10] [11].

One interesting paper that combines the technology acceptance model with the didactic cycle was published recently in 2014 by Claus-Peter H. Ernst et. al. [12] and they state that despite the advantages of e-learning platform they are usually underutilized. The attitude towards learning whole learner's corresponding whole life attitudes. The authors seek to identify potential influence factors of students toward attitude using e-learning technologies.

There are also two papers that need to be mentioned as related work because this paper is a follow up. The first study regarding the usability of Tesys e-Learning platform [13] tackled the problem of the student's interface but used a small group of students which saw for the first time the e-Learning platform. The results revealed some bugs and issues which were already addressed but due to the small number of participants there was not too much knowledge gained regarding the usability and no statistical methods could be applied. The second one [14] extended the first study but this time using participants that already used the platform. The second study was larger and we integrated the questionnaire in the platform using google docs, so the students could access the questionnaire for big amount of time. We also used in that paper Wilcoxon method to separate groups and in the case of students we got significant

differences between questions.

- questions for group separation (GS).

Regarding the ease of use we proposed fourteen questions that were meant to reveal if there are any problems, these questions were meant to be validated by the control questions (6): 4 for the groups of questions, one for overall validation and one for explanations regarding the help section that also validates the previous question because if the answer was low but no suggestions were made there is a certain inconsistency. If there was inconsistency between

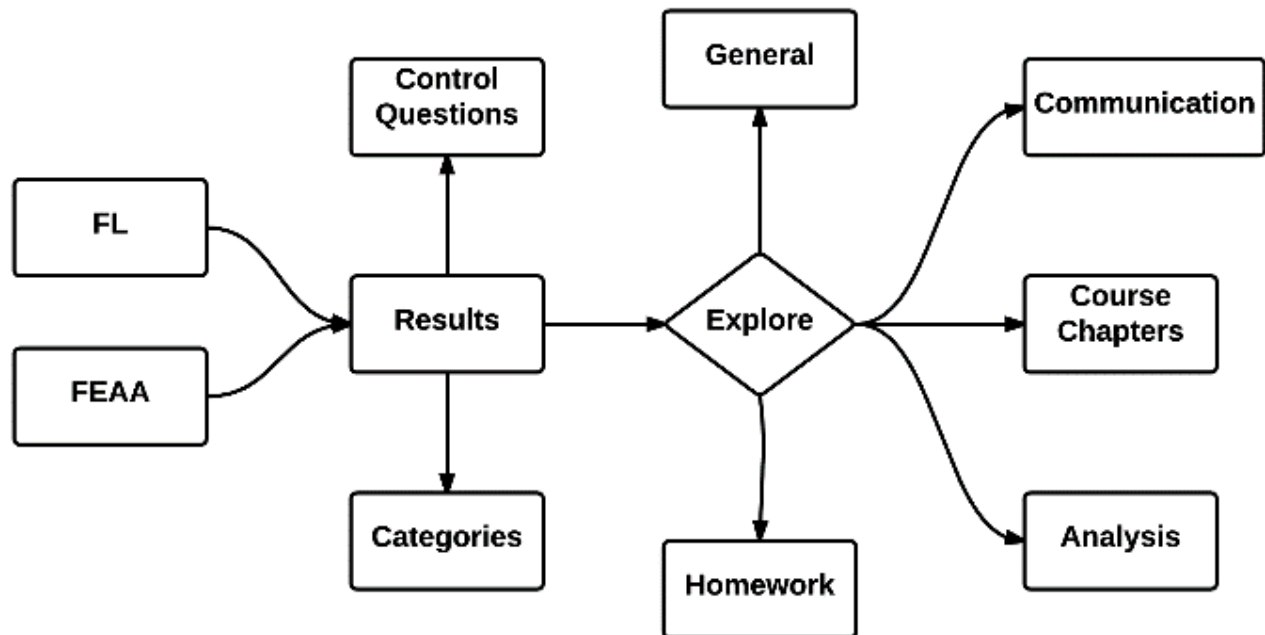


Figure 1. Design Flow

**STUDY SETUP**

Figure 1 presents the overall flow for the study. We gathered results from FL and FEAA, and the results contained the studied questions but also control and categories question on which we perform the statistical analysis. The main questions explored 5 principal categories: homework module, course module, communication module, the students’ analysis module, and one section that aimed to evaluate the main interface.

On this study, 34 professors completed the questionnaire, 18 from FL and 16 from FEAA. The questionnaire has twenty-four questions divided in three main categories:

- questions regarding the platform’s ease of use (EU),
- control questions (C)

the control question and the group of questions validated by it the questionnaire was disregarded. Four questions were used for group separation and based on them we are able also to get an intuition regarding the level of experience of the responder and how often do they use the platform.

No	Question	Type
1	How easy do you find the courses that you teach?	EU
2	Have you ever used the manage function of a discipline?	C
3	Organizing chapters, homework, links on the course management page is intuitive?	EU
4	From 1 to 5, how easy was the process of adding a new chapter?	EU
5	From 1 to 5, how easy is to setup a new question for the course chapters?	EU
6	Until now, have you ever used the	C

	homework management function.	
7	From 1 to 5, how easy is to add a new homework?	EU
8	From 1 to 5, how easy is to give grades to homework?	EU
9	Adding external references (links) for the disciplines is easy to use?	EU
10	Did you found fast the external references (links)?	EU
11	Have you ever used the analysis module from the disciplines?	C
12	The data regarding the students was enough?	C
13	How intuitive is the data regarding the students presented in the analysis section?	EU
14	Have you ever used the Communication section?	C
15	From 1 to 5 how easy did you used the section “communication with students”?	EU
16	From 1 to 5 how easy did you used the section “communication with secretaries”?	EU
17	From 1 to 5 how easy did you used the section “communication with professors”?	EU
18	The “Help” section offers enough data?	EU
19	If you choose “no”, please describe what it should contain?	C
20	How many disciplines do you teach at distance learning?	GS
21	At what year of study do you teach?	GS
22	How many years of experience do you have at distance learning?	GS
23	How many hours a week do you use the e-Learning platform?	GS
24	What overall grade do you give to the platform?	C

Table 1. Questions and groups

Table 1. presents the questionnaire setup. On the first column, we have the question number, the second column has the questions and the last describes the type of the question. In the paper, the number of questions will be further addressed.

### EXPERIMENTAL RESULTS

In order to explore the perceived ease of use, we considered

two of the group separation questions and we explored if the tested factors influence somehow the perceived ease of use. We disregarded for the study question twenty, and twenty-one because the lack of information, most of the professors that completed the questionnaire were teaching at almost all years, and the number of disciplines was from one to three for 94% of the professors from FL and 100% from FEAA.

Count of "How many years passed since you started using Tesys?"

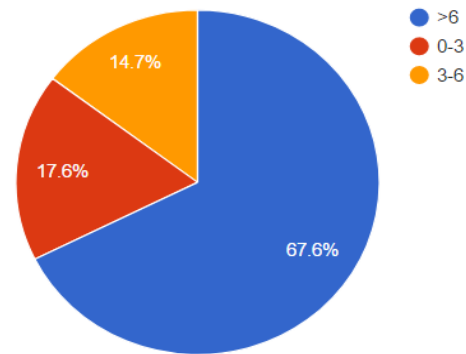


Figure 1. Years spent on platform

Figure 1 presents the number of years since the professor's that attended this study started to use Tesys. This answer somehow motivates the small number of hours spent weekly because if they already have a very good experience with the platform there is no need to spend many hours on the platform in order to manage courses, questions etc.

In order to have significant groups we decided to have a threshold of 6 years and we merged the two groups from zero to three and from three to six years and we obtained one which is referred as less than six years in the next table.

QID	p value	W statistic	Average for <6 years	Average for >6 years
1	0.585	115	4.739	4.181
3	0.736	117.5	4.260	3.818
4	0.864	122	4.565	4.090
5	0.675	137.5	4.086	3.909
7	0.697	136.5	4.347	4.090
8	0.346	102	4.173	3.454
9	0.625	113.5	4.130	3.545
10	0.431	106	4.043	3.545



13	0.788	119	3.913	3.545
15	0.860	121.5	4.130	3.727
16	0.381	149.5	3.782	3.909
17	0.877	131	3.826	3.636
18	0.741	135.5	3.869	3.818
24	0.798	119.5	4.217	3.818

Table 2. Wilcoxon results on number of years spent on the platform

Table 2 presents the results obtained after separation of the data in two groups based on the number of years on which the professor that answered performed his activities on the e-Learning platforms.

The first column from the table represents the question's number, the second the w statistic, then, the third is regarding the p value and the last two columns represent the mean values for the chosen categories. Based on the data collected we defined two groups, professors that have an experience of over six years in the Tesys e-learning platform, and professors that have an experience below six years. We use w statistic and p value to get an intuition regarding how significant is the difference between the two categories.

Based on the results we collected we can say that the number of years spent on the e-learning platform do not influence the grade that was given to the questions.

Count of "How many hours do you use the platform weekly?"

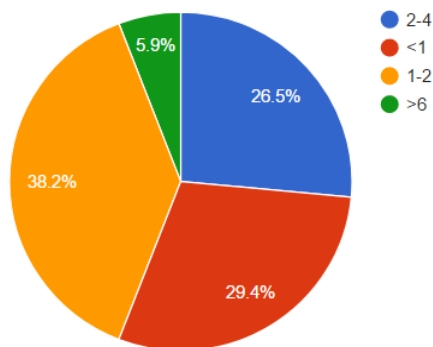


Figure 2. No. of hours spent on platform

Figure 2 presents the distribution of the number of hours spent on the platform weekly. From the figure, we can conclude that professors used the platform very low, most of the professor's spend less than two hours per week on the platform. In the next table, we present the results for two groups, we merged that had less than one hour with the

one that spent one to two hours and then the other two ones obtaining two more comparable groups on which we could use Wilcoxon statistic.

QID	p	W statistic	Average grade for <2 hours	Average grade for >2 hours
1	0.961	133.5	4.666	4.5
3	0.588	146.5	4.083	4.136
4	0.136	168	4.333	4.454
5	0.906	128.5	4.083	4
7	0.602	118.5	4.416	4.181
8	0.630	119	4.166	3.818
9	0.632	119	4.166	3.818
10	0.551	116	4.083	3.772
13	0.638	119	3.916	3.727
15	0.744	141	3.916	4.045
16	0.924	129	3.833	3.818
17	0.954	134	3.666	3.818
18	0.362	156.5	3.583	4
24	0.953	130	4.166	4.045

Table 3. Wilcoxon results on number of hours per week spent on the platform

Table 3 presents the results after exploring if the number of hours spent on the platform offers significant difference between groups. The chosen threshold tested was of two hours spent in one week.

The first column from the table represents the question's number, the second the w statistic, then, the third is regarding the p value and the last two columns represent the mean values for the chosen categories.

Analysing the results, we can say that also in this case the number of hours do not influence the perceived ease of use regarding the e-learning platform.

### CONCLUSIONS

In this paper is presented an exploratory analysis in order to see how Tesys e-Learning platform is accepted among professors that teach in this learning environment. The results highlighted that no matter what time do they spend weekly on the platform or their experience, the professors accept with no significant differences the way Tesys is built. The overall grade of the platform is above four, so on a scale from one to five, the overall grade indicates a good technology acceptance.

The limitation of this study is represented by the dwindling number of professors that answered our questionnaires and the small number of group separation questions. Even

though in the questionnaire there are more group separation questions, the groups distribution was not good enough to explore the data. Increasing the audience of the study may provide more data and better group separation and this would have been possible if the study would have taken more time.

The main conclusion that can be taken from this paper is that the experience and the number of hours spent on the platform do not influence the way professors accept Tesys e-Learning platform and on a scale from one to five, the mean overall grade is above four. There were a few issues highlighted in the communication module which is already under the solving procedure.

#### **ACKNOWLEDGMENTS**

This work was supported by QFORIT Programme, University of Craiova, 2017.

#### **REFERENCES**

1. Burdescu, D. D., & Mihaescu, M. C. (2006). TESYS: e-Learning Application Built on a Web Platform. In ICE-B (pp. 315-318).
2. Costabile, M. F., De Marsico, M., Lanzilotti, R., Plantamura, V. L., & Roselli, T. (2005, January). On the usability evaluation of e-learning applications. In System Sciences, 2005. HICSS'05. Proceedings of the 38th Annual Hawaii International Conference on (pp. 6b-6b). IEEE.
3. Ardito, C., Costabile, M. F., De Marsico, M., Lanzilotti, R., Levialdi, S., Roselli, T., & Rossano, V. (2006). An approach to usability evaluation of e-learning applications. *Universal access in the information society*, 4(3), 270-283.
4. Goldberg, J. H., & Kotval, X. P. (1999). Computer interface evaluation using eye movements: methods and constructs. *International Journal of Industrial Ergonomics*, 24(6), 631-645.
5. Bowman, D. A., Gabbard, J. L., & Hix, D. (2002). A survey of usability evaluation in virtual environments: classification and comparison of methods. *Presence: Teleoperators and Virtual Environments*, 11(4), 404-424.
6. CHIN, P., DIEHL, A., & NORMAN, K. (1988). Development of a tool measuring user satisfaction of the human-computer interface. Disponível por WWW em <http://lap.umd.edu/lapfolder/papers/cdn.html> (jun. 1998), 92.
7. Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319-340.
8. Al-Adwan, A., Al-Adwan, A., & Smedley, J. (2013). Exploring students acceptance of e-learning using Technology Acceptance Model in Jordanian universities. *International Journal of Education and Development using Information and Communication Technology*, 9(2), 4.
9. Bland, J. M., & Altman, D. G. (1997). Statistics notes: Cronbach's alpha. *Bmj*, 314(7080), 572.
10. Park, S. Y. (2009). An analysis of the technology acceptance model in understanding university students' behavioral intention to use e-learning. *Educational technology & society*, 12(3), 150-162.
11. Masrom, M. (2007). Technology acceptance model and e-learning. *Technology*, 21(24), 81.
12. Ernst, C. P. H., Wedel, K., & Rothlauf, F. (2014). Students' Acceptance of E-Learning Technologies: Combining the Technology Acceptance Model with the Didactic Circle.
13. Popescu, P. Ș., Mihaescu, M. C., Mocanu, M., & Ionașcu, C. (2016). Evaluation of Tesys e-Learning Platform's Interface. In ROCHI-INTERNATIONAL CONFERENCE ON HUMAN-COMPUTER INTERACTION (p. 86).
14. Mihaescu, M. C., Popescu, P. Ș., Ionașcu, C. (2017). Questionnaire Analysis for Improvement of Student's Interaction in Tesys e-Learning Platform in Romanian *Journal of Human - Computer Interaction*, Vol.10, No.1.

# Learn-o-Bot: Educational Application Using Augmented Reality and Sphero SPRK+

**Andreea Camelia Tripon**

Computer Science Department,  
Technical University of Cluj-Napoca  
Cluj-Napoca, Romania  
camelia.tripon1@gmail.com

**Teodor Stefanut**

Computer Science Department,  
Technical University of Cluj-Napoca  
Cluj-Napoca, Romania  
teodor.stefanut@cs.utcluj.ro

## ABSTRACT

This paper presents a mobile application with educational purposes designed for kindergarten children which is called Learn-o-Bot. The application uses augmented reality and the Sphero SPRK+ robot to create an appropriate environment for learning.

The purpose of Learn-o-Bot is to provide preschoolers with means to acquire basic knowledge through a fun game which is easy to understand and use. The game consists of five levels with gradually increasing difficulty. Each level's scene is augmented with appropriate virtual objects and the child should control the robot to select the right objects and reach level's goal.

Finally, this application opens a new perspective to different learning methods, using robots, mobile devices and virtual learning materials.

## Author Keywords

Augmented Reality; Sphero; Education; Mobile applications.

## ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

## General Terms

Human Factors; Design.

## INTRODUCTION

Kindergarten children have a broad capacity to accumulate knowledge, but they also lose their focus very quickly. Therefore, they need an attractive learning environment, which is able to keep them interested and to provide them with the proper knowledge for their age. We propose this augmented reality game in which the Sphero SPRK+ robot is a bridge between the real world and a virtual one, designed towards learning purposes.

## Traditional Educational Methods

The classical educational methods are still very popular in kindergartens and consist in involving the child in activities like drawing, singing or solving simple tasks. The main disadvantage is that children get distracted before finishing all the tasks and that these methods require an adult, an instructor to assign activities and supervise children's progress.

## Technology in Education

Using technology in kindergartens can bring many benefits to the learning process, making it both fun and interesting. For example, children can listen to an alphabet song, watch movies or cartoons, use tablets or smartphones to play educational games.

Children also love toys, and with the increased development of technology, they prefer toys like tele-guided cars or robots. Kids these days already know how to use a smartphone [1], a laptop and how to control a tele-guided car, therefore a ball controlled by the phone is a good choice to draw their attention

## Augmented Reality and Education

Augmented reality (AR) allows users to view virtual objects on top of the physical world observed through the mobile device's camera.

In education, AR would bring numerous advantages and there are infinite ways of using it. It can make every classical learning activity way more interesting. For example, instead of hearing stories and viewing pictures of wild animals, the child can see them in 3D with proper sounds and animations. Math can become funnier if, for example, a card with a number is augmented with the same number of animals as in the game Math Alive [2].

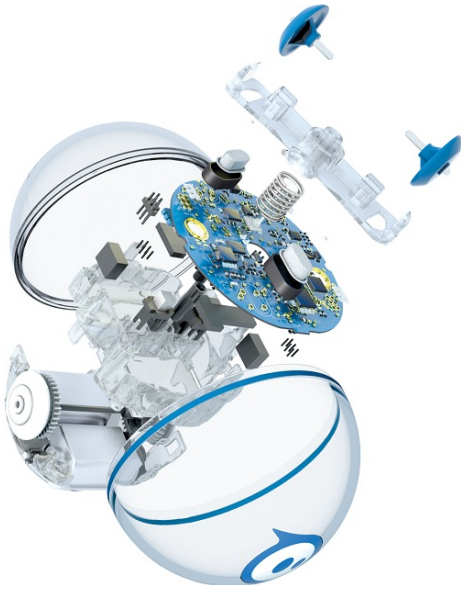
In the following sections, we will present the Sphero SPRK+ robot, the project components, algorithms and the game from a user perspective (each level of the game). The paper will end with the conclusions and future development.

## SPHERO SPRK+

Sphero SPRK+ [3] is a ball-shaped robot designed for educational purposes which can be controlled by a mobile device or computer. The robot can move in any direction with the selected speed and has 2 RGB LEDs, feature which make this toy even more attractive for kids. One important design feature is that Sphero's shell is transparent and kids can actually see its electronics (see Figure 1).

## Hardware specifications

The robot is equipped with accelerometer and gyroscope sensors, which are useful to determine collisions, measure velocity or determine robot's orientation.



**Figure 1. Sphero SPRK+ components [4]**

The communication between Sphero and the controlling device is made through Bluetooth. Sphero SPRK+ uses Bluetooth 4.0, known as Bluetooth Smart or BLE. This feature is the main upgrade from the previous version, Sphero SPRK, which required additional steps in order to establish a connection. The robot can be controlled from a distance up to 30 m.

*Software specifications*

Sphero’s creators offer support for different platforms as Android, iOS, Windows, Unity, through a Software Development Kit (SDK). These libraries are used to send proper commands to the robot. Sphero understands commands such as: driving in a direction with a selected speed, changing the color of the LEDs or changing the orientation. The SDK also handles connection lifecycle and can receive asynchronous information from the sensors.

**RELATED WORK**

The main features of the current application imply using Sphero and AR for educational purposes. Since Sphero itself was designed for learning, there is a large variety of related applications which are used in this scope.

The most significant is the Lightning Lab [5] application which was designed by the same company as the robot. The application contains a full set of instructions to control the robot and read its state. It also has a feature that allows users to develop small programs that can be executed by the robot. These programs can help users understand basic programming principles as conditional blocks or loops.

Another example where Sphero is used as an educational mean is presented in article [6]. Since Sphero offers free SDKs to allow developers design their own programs, the authors are using this particularity to teach students basic Object Orientated Programming (OOP) principles. The paper presents a study where university students, with no

background in programming, can learn basic C# and programming principles by designing a program which controls the Sphero robot. The students claim that this learning method is more interesting and effective because they can see the practical use of programming by controlling Sphero’s moves.

Sphero and AR meet in another application made by the creators of Sphero: The Rolling Dead [7]. The gameplay consists in using Sphero as the main character in order to defeat the zombies that are spawning in an augmented environment.

SMAUG is a collaborative AR game that also uses Sphero as the main character. The game is presented in article [8], where the authors show that Sphero’s control can be shared to complete more complex tasks. The environment is augmented with objects like diamonds, dragons or treasures and Sphero has to collect or avoid these virtual objects to complete the level’s goal.

Quiver [9] is a game which uses augmented reality to create living drawings. It is designed for children to explore their creative aptitudes. The game consists in printing and coloring the drawings and then using the app to see the drawing coming to life. The interesting part is that the colors used in the real drawing are also used for the augmentation, giving the artist a special sense of ownership. Also the characters are animated and the user can interact with them by tapping on the screen and selecting different actions. The purpose of this game is to make coloring more fun, interesting and to allow the child to interact with the drawings.

Zoo-AR [10] is another educational augmented reality game where children can use custom pictures with the animals to see them in 3D. Although it does not use Sphero, the application is a good educational app example because it is a fun way to learn about animals and users can interact with them or take screenshots. It is useful to see how the animals would behave in real life, without actually being in their natural habitat.

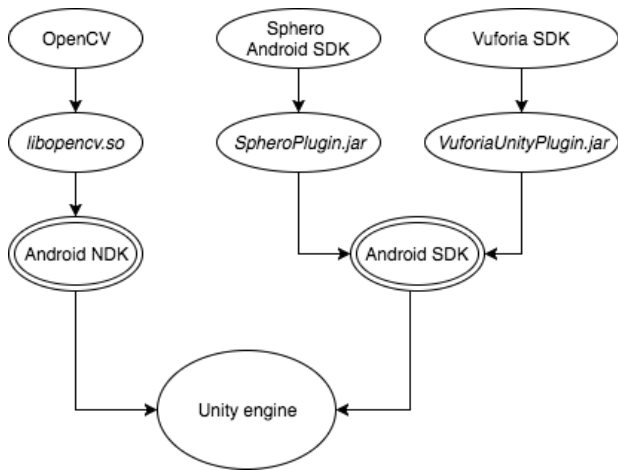
**PROJECT STRUCTURE AND DEPENDENCIES**

The application is made using Unity development environment. Unity permits to create cross-platform games, but this application is currently available only for Android devices.

*Software Dependecies*

Figure 2 shows a diagram of all the libraries and dependencies used to create this application. These libraries are presented as follows:

1. Sphero Android SDK [11] – a library which is used to send commands to the robot. Since Bluetooth 4.0 is a newer version, the open source SDK is only in beta version and is not available for Unity. To overcome this obstacle, we created an Android plugin for Unity which acts as a bridge between Unity procedures and Java methods provided by the



**Figure 2. Dependencies diagram**

Sphero SDK. The plugin supports simple instructions like connecting to the robot, sending commands to move or to change colors and calibrating.

2. Vuforia [12] – a library used for AR support for Android, iOS or Unity. It is used directly as a plugin for Unity and provides scripts and prefabs (objects already built in the library that include the necessary properties and scripts) to recognize markers and show augmented objects on top of the target area. The markers supported by Vuforia are: image, frame, object (cylinder or custom object that is scanned and loaded into the app). For this application, we used a simple image target evaluated with 5 stars on Vuforia website. All the scenes that use augmented reality use a prefab called ARCamera and an ImageTarget.

3. OpenCV [13] – is a library for image processing. It is used to detect the Sphero object in the image captured by the camera. Because it is not available for Unity, we wrote a small program in C++ which is then compiled in a shared object (.so) and called using NDK directives. The shared object library has dependencies to the open source OpenCV library and contains the processing method necessary to detect the robot.

#### Application scenes

Unity uses scenes to distinct the “pages” of the game. In the current application, we have 5 scenes, one for each level, and four others which provide a complete user experience in the game:

- *MainMenu* – this scene contains only UI objects: three buttons that redirects the user to the desired action: *Play*, *Drive sphero*, *Help*
- *DriveSphero* – this scene is dedicated entirely to controlling Sphero. It does not contain any AR layer because it acts as a training room to get used to the commands for directions, speed or calibration process.
- *HelpScene* – shows all the information about the game, connection to Sphero or how to use the app.

- *LevelsMap* – this scene contains UI elements that allows the user to select one of the five levels available. The screen also displays the rank (1, 2 or 3 stars) and level’s availability (locked/unlocked).

## DEVELOPMENT SOLUTIONS

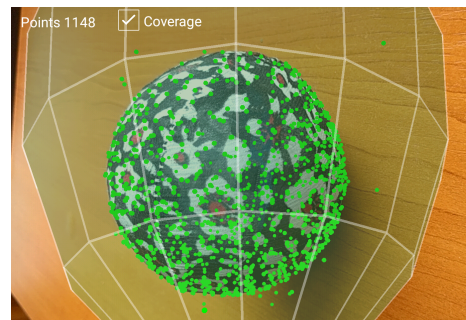
### Sphero Tracking Algorithm

Sphero is the main character of the game and we must always know its position in the 3D scene. To do that we tried two different approaches: using Vuforia library and using a custom algorithm to detect light sources.

#### Tracking Sphero with Vuforia

Vuforia offers support to track 3D objects by scanning all the faces of the object and uploading the scan into the developer portal. The scanning process finds feature points on the object which will be later recognized and tracked from the video feed.

The problem is that Vuforia can track objects that have many edges and contrasting areas, which is exactly the opposite of Sphero’s shape and colors: no edges and only shiny surfaces. To overcome at least one of these obstacles, we decided to wrap the ball with a material that has a lot of features. A picture of the used fabric was first uploaded to Vuforia developer portal and it was evaluated with 5 stars rating. We can see the wrapped Sphero and the coverage from scanning in Figure 3. As it is observed in the picture, Vuforia found 1148 points of interest that will be used to track the ball.



**Figure 3. Wrapped Sphero and interest points**

Although the scanning algorithm found many points of interest, the tracking results were not very satisfactory. The ball was detected only from small distances and if the room was well-lit. The tracking was quite unreliable if the ball was moving or the distance was more than 50 cm. Because the results from this approach were not good enough, we decided to use another method for tracking.

#### Tracking Sphero as a Light Source

Since Sphero’s characteristics were not easily detectable with the previous method, we decided to use the light source and the spherical form for tracking Sphero. Using the image captured from the video feed, in each frame, the following algorithm is applied:

- transform image to grayscale
- binarize image using a threshold of 245



- noise reduction by applying the morphological operation “opening” with a 7x7 kernel
- find the contours from the remaining data in the image. The center of the contour is sent to Unity, where an algorithm is applied to transform the 2D point in a 3D object.

These operations are written in C++ and use the OpenCV library to process the image. The result of this algorithm is shown in Figure 4. The disadvantage of this method comes when more than one light source can be identified in the image (see Figure 4b.)

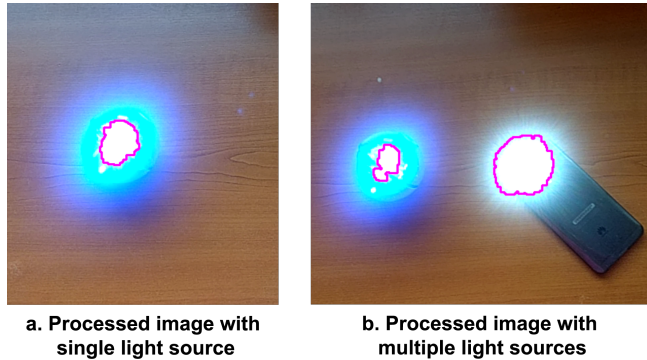


Figure 4. Processed image with OpenCV

#### Transform 2D point to 3D object

Because we want to have Sphero as a virtual object in the scene, we need to translate the light source center point which is in screen coordinates (2D) into scene coordinates. The virtual scene is positioned based on a custom image marker detected by Vuforia. This marker and the robot should be on the same plane, meaning that we can determine Sphero’s 3D position by intersecting the ray starting from the screen point with the plane determined by the marker as observed in Figure 5.

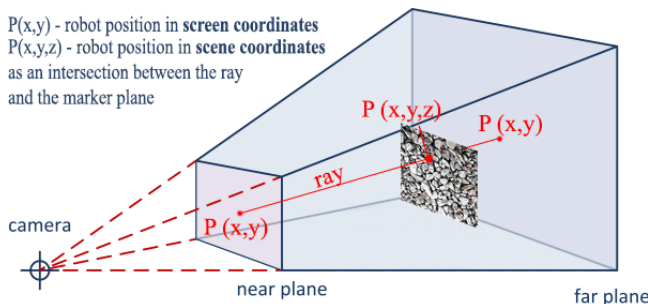


Figure 5. Screen position to 3D position

After knowing the 3D position of Sphero’s center, all we have to do is create a 3D object with Sphero’s dimensions and move it to the computed point in scene. The algorithm performs well on medium performance devices and the position is updated in each frame following the light source.

#### Sphero Auto-Calibration Algorithm

Being a spherical object, we do not know which way Sphero considers to be the “forward” direction. To be sure that

Sphero is moving according to our joystick input, before using it, we should calibrate the robot. In other words, we should teach the robot which way is forward for us. This step consists in aligning the robot backlight in front of the user, by rotating the robot.

Although calibration is not difficult, children would prefer to just dive into the game, without any additional steps for the setup. Also, they can do mistakes in calibration which will result in sending wrong directions to Sphero. For these reasons, we decided to develop an auto-calibration feature. This consists in automatically aligning the robot to user forward direction, when the game is started.

The algorithm is represented in Figure 6 and has the following steps:

- choose a position  $S_{target}$  in front of Sphero considering camera position. The distance between Sphero and position  $S_{target}$  is  $d$
- move Sphero with the direction “forward” (0 degrees) until  $d$  is the distance between Sphero’s initial position  $S_{init}$  and Sphero’s current position  $S_{curr}$
- compute the angle between the vectors determined by  $S_{init}S_{target}$  and  $S_{init}S_{curr}$  and adjust Sphero’s rotation with this angle.

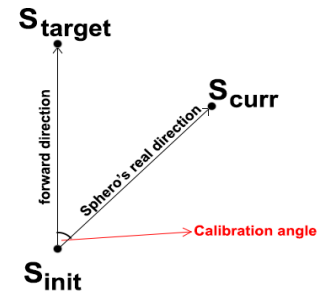


Figure 6. Calibration algorithm representation

#### GAME STORIES

As it we previously mentioned, the game is divided in five levels with progressive difficulty. Each level will teach the child one different concept appropriate to his age (numbers, geometrical figures, animals, etc). The user can play a level, only if he/she completed all the previous levels.

#### Common Elements in Levels

A level is represented as a Unity scene. Sphero is represented in all scenes by a translucent sphere and its position is updated according to the light source center. The collision detection is performed in every frame and is based on the virtual representation of Sphero and the virtual objects placed on the scene. A collision between virtual Sphero and another object in scene is followed by visual and audio indicators: Sphero will have green lights when the object is correct and red lights when a wrong object is selected and a proper sound will be played in both cases.

#### Star rating

To evaluate a level, we use the number of mistakes and the time spent to achieve the level’s goal. Each level has



different thresholds for these parameters, depending on the number of elements and the complexity of the scene. The performance on each level will be ranked with a rating from 1 to 3 stars. Three stars rating can be achieved only if the user did not make any mistakes like selecting a wrong object and if he/she completed the level in a reasonable time.

### Level UI

Every level has one UI layer with the following elements:

- *joystick* – it controls the robot and accepts moves in any direction from 0 to 360 degrees
- *calibration button* – initiates the calibration of the robot
- *elapsed time* – shows the time spent in the current level
- *mistakes count* – shows the number of wrong selected objects
- *reset button* – resets level’s progress, elapsed time and number of mistakes
- *help button* – replays instructions for the level
- *back button* – goes back to level selection scene
- *start level dialog* – shows a short description of the level and also plays an audio with the instructions
- *end level dialog* – shows the score of the level (1, 2 or 3 stars), the time and the number of mistakes.

### Levels Description

#### Level 1 – Colors and 3D shapes

This level will help children to accumulate basic knowledge about colors and shapes. Although this kind of exercise is commonly used in kindergartens, this method is funnier and more interesting. The scene consists of 7 geometrical figures like cube, sphere, cylinder which are colored in 4 base colors (see Figure 7). The gameplay consists in selecting the correct objects as requested in a message at level start and after each correctly selected object. There are no restrictions regarding the path to the correct object, but Sphero should not collide with a different object because it will be considered a mistake.



Figure 7. Level 1 - Colors and 3D shapes

#### Level 2 – Learning numbers

Since numbers are one of the most important concepts that should be learned by children, the second level proposes a fun way to learn counting. The scene shows the numbers from 1 to 9 in a random order (see Figure 8) and the child must select all numbers in ascending order. If mistakes are

made, these will affect the final score, but the user can continue selecting another number without restarting the level.



Figure 8. Level 2 - Learning numbers

#### Level 3 – Wild Animals

One important part in child’s development at a preschool age is to learn the animals and their habitat. Because some of these animals could be dangerous, this application proposes another way than seeing them in real life. 3D projections of animals will appear in the scene (see Figure 9) and the child can interact with them by tapping the screen. For now, the interaction is basic: play the sound of the animal and execute an animation, but more complex interactions will be added.

To pass this level, user has to select each animal in the correct order. The order will be displayed step by step: first message is shown at level start and the following are displayed after each correct selection.

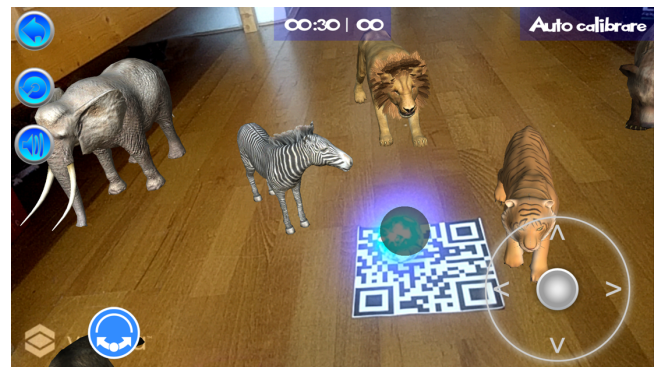


Figure 9. Level 3 - Wild animals

#### Level 4 – Drawing dot-by-dot

Drawing is another common activity in kindergarten. For this application, we included the dot-by-dot drawing technique which is both fun and educational. Dot-by-dot drawing implies showing a pattern and the control points of that drawing. The child should connect the points in the correct order to obtain the final figure. For this application, we included 2 simple patterns for now: house and star. The control points are shown as letters (see Figure 10). Therefore, while playing this level, the child will learn the first letters of the alphabet and to draw simple figures.

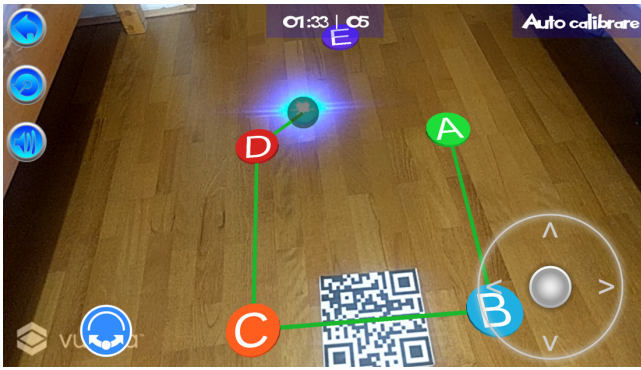


Figure 10. Level 4 - Drawing dot-by-dot

#### Level 5 – Collect Chests

This level is a little different than the others. It does not teach the child any new concepts but rather it tests his abilities to control and to think ahead his moves.

The scene is composed of 3 moving obstacles and 3 chests (see Figure 11). To complete the level, the child must collect all the chests but without touching the obstacles. The first two obstacles are boxes that are moving with different speed on Ox axis. To collect the first two chests, user should observe the course and the speed of the boxes and carefully control Sphero to avoid those obstacles. The third obstacle is a moving rock that is rotating and protecting the final chest. The child must collect the final chest when the rock passed, but he should be careful not to get hit when it rotates back.

When Sphero is hit by one of the obstacles, the level is reset and Sphero is sent back to the starting point. No mistakes are permitted to reach level's goal.

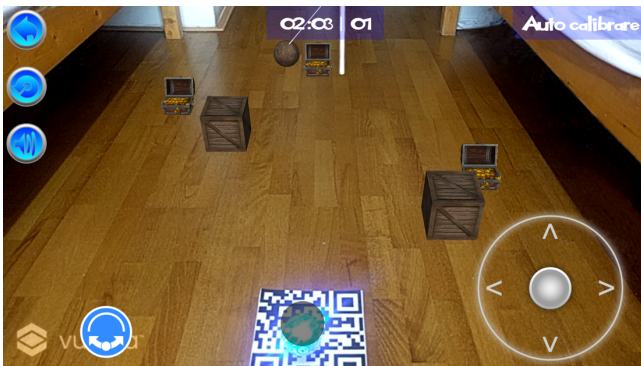


Figure 11. Level 5 - Collect chests

#### CONCLUSIONS

As a first conclusion, the proposed application can be considered an innovation for supporting preschoolers teaching process. Learn-o-Bot can be used both at home or in kindergartens and it does not require a supervisor to explain the task to the child. The messages are clear and always appear in two forms: written and audio, making the communication process easier.

The actual learning process happens when the kid controls the robot into achieving level's goal. If the child completes a level, it is considered that he accumulated the necessary information from that level and can pass to the next one. The star rating given at level end acts like a reward and a grade for finishing the level. To not disappoint kids, at least one star is given just for finishing the level. This rating will stimulate him/her to try again and receive the maximum number of stars. This means that a small rating will most likely make the child replay the level and get better rating, while a three-star rating means that the concepts are learned and understood correctly, without mistakes and in a good response time.

In current state, the application offers support only for a small area of knowledge that is necessary for children. To become a better educational application, the first improvement will be to add more levels in order to cover as much information as it is possible. Also, for now, the speed of the robot is fixed and cannot be changed from within the app. As an improvement idea, a good feature would be to allow the users to control the speed (for example in games with car driving, a slide from bottom to top means accelerating). And of course, an important step would be to improve Sphero's tracking algorithm.

#### REFERENCES

1. Rosie Flewitt, David Messer, Natalia Kucircova, "New directions for early literacy: The iPad", Journal of Early Childhood Literacy
2. Sphero SPRK+: <http://www.sphero.com/sprk-plus>
3. Math alive website: <http://alivestudiosco.com/math-alive-kit/>
4. Sphero's components image source: <https://www.maximisetechology.com.au/product/sphero-sprk-plus/>, available on 25 May 2017
5. Lighting Lab, <http://sprk.sphero.com/>
6. J. M. Rodriguez Corral, A Morgado-Estevez, D. Molina Cabrena, F. Perez-Pena, C. A. Amaya Rodriguez and A. Civit Balcells, *Application of Robot Programming to the Teaching of Object-Oriented Computer Languages*, International Journal of Engineering Education, 2016
7. The Rolling Dead: <http://blog.sphero.com/blog/spheros-the-rolling-dead-now-available-on-ios-and-android/>
8. M. Panzariu and A. Iftene, *Sphero - Multiplayer Augmented Game (SMAUG)*, RoCHI 2016
9. Quiver website: <http://www.quivervision.com/>
10. Zoo-AR website: <http://zoo-ar.com>
11. Sphero Android SDK source: <https://github.com/orbotix/Sphero-Android-SDK>
12. Vuforia, <http://vuforia.com/>
13. OpenCV documentation, <http://docs.opencv.org/>

# Overview of virtual worlds usage for student engagement

**Ing. Stefan-Lucian Bucur**  
POLITEHNICA University of  
Bucharest  
Splaiul Independenței 313,  
București 060042, Romania  
slbucur@gmail.com

**Prof. Dr. Ing. Florica  
Moldoveanu**  
POLITEHNICA University of  
Bucharest  
Splaiul Independenței 313,  
București 060042, Romania  
florica.moldoveanu@cs.pub.ro

**Prof. Dr. Ing. Alin  
Moldoveanu**  
POLITEHNICA University of  
Bucharest  
Splaiul Independenței 313,  
București 060042, Romania  
alin.moldoveanu@cs.pub.ro

## ABSTRACT

This paper presents an overview of 3D virtual world systems used by universities around the world to increase student engagement. Using publicly available information from the universities websites, we classified the virtual world systems they have deployed according to the type of technology used, availability and level of immersion. Using this classification, we provide a high-level view of the currently used systems and the expectations set for them by their target audience, students and prospective students. The purpose of this paper is to analyse the state of the art in the field of virtual world systems used in universities, as a start point for further improvements in this field. We also present a set of emerging technologies used in this context: augmented world through personal smartphones and virtual reality helmets.

## Author Keywords

Virtual worlds; education; student engagement

## ACM Classification Keywords

H.5.1. Information interfaces and presentation: Multimedia Information Systems.

## General Terms

Experimentation; Human Factors; Design;

## INTRODUCTION

Over the last decades, 3D virtual world systems have evolved from crude simulations to full immersive experiences. With their increased popularity among young people and the appearance of affordable devices (personal computers and smartphones) that can render them, universities around the world are using 3D world simulations to increase student engagement.

Virtual world systems are 3D simulated worlds rendered on a screen using a computing device, video games representing the most common example. They range from simple, where the user can see panorama style pictures from fixed locations but cannot move through the world, to complex, in which the user can navigate through a fully simulated 3D environment.

The universities most commonly use these simulations for three purposes:

- **presentational**, by showing the campus and associated facilities to prospective students, for example in 3D virtual tours;
- **social**, as an online hub for students to communicate, for example multiplayer campus simulations where the students can post messages;
- **educational**, by simulating learning materials and experiments, for example interactive 3D courses;

## History

The first virtual campus simulator used by a university was the one developed by the Nanyang Technological University in Singapore (NTU) since 1997 [1]. The system is no longer maintained, but at its appearance, it featured a full 3D simulation of the university campus and also interactive lessons for 3D modelling.

This system appears to be ahead of its time, as no other similar offerings have been developed until 2009 at other Asian universities [2], [3] and [4].

Because of high costs of development and maintenance, these types of fully immersive systems did not become widespread. As such, most public facing virtual world systems currently deployed by universities are 3D virtual tours available online [5], [6] and [7]. These tours are limited in scope and have lower production costs, as they only show 360-degree pictures and videos of the campus from fixed locations.

With the improvement of frameworks for developing 3D applications like Unity 5 [8] and Unreal Engine 4 [9] that allow the creation of fully immersive environments at much lower production costs, it is possible that improved campus simulations will be available in the near future.

In contrast to virtual campuses, 3D courses had a linear progress. Currently this type of courses are widely available and cover a variety of subjects, like DNA simulation [10] and surgery [11]. In addition, dedicated platforms exist to facilitate the development of 3D virtual classrooms [12].



## CLASIFICACION OF VIRTUAL WORLD SYSTEMS

Considering student engagement, virtual world systems can be classified based on three characteristics:

- Used technology
- Availability
- Immersion

### Type of technology

Developers have two options for simulating 3D worlds: static images/videos or interactive simulated 3D environments.

**Static images and videos** are made using a technique known as 360-degree spherical photography or virtual reality photography. Until recently, this procedure required expensive panoramic cameras, but with the appearance of digital cameras, a process known as “stitching” can be used, which uses multiple rectangular pictures and algorithmically unifies them to a single panorama [13].



**Figure 1. 360-degree spherical photo projected on a flat surface**

Users can view the produced 360-degree photo using any device, through a browser or a specialized application. One popular service that uses this technique is the Street View from Google Maps [31].

**Interactive 3D environments** are most commonly associated with video games. They use 3D models, textures and animations to render a virtual world on the user’s screen.



**Figure 2. Rendered 3D environment [16]**

Creating 3D environments is usually more expensive than 360-degree photos or videos, as a team of artists has to create every required asset manually.

Through a technique known as photogrammetry, real objects can be used to generate 3D models and textures. This is especially useful for simulating places from the real world which are typically very difficult (time consuming and expensive) to model.

### Availability

Availability refers to how easily users can access the virtual world and in general refers to the hardware, software and time requirements of the product.

From a **hardware** perspective, a system that can run on more devices is considered more available. 360-degree photographs can be rendered on any device and have low requirements. Interactive 3D worlds, on the other hand, need resources proportional to their visual fidelity. Displaying a realistic 3D environment needs much more computing power than displaying a single photograph.

From a **software** perspective, a system is considered widely available if it does not need the installation of additional software packages.

Web applications are a type of widely available software. A user only needs to access an URL in a browser and is instantly able to use their functionality.

Because of their low computing requirements, 360-degree pictures are readily accessible online.

Virtual 3D environments can also be made available online, but rendering them through a browser increases their already high requirements, as the browser abstracts underlying hardware and makes it harder for applications to use native instructions.

In recent years, this problem was alleviated through the introduction and improvement of WebGL [14], which allows applications to use the graphical pipeline available on hardware systems in direct manner. Also in this direction, WebAssembly [15] will help increase the performance of web applications with high computing requirements, but this is still a work in progress.

In general, rendering 3D environments requires installing software on the user’s device. This process can be facilitated through software installers, like application stores.

### Immersion

Immersion refers to the deep mental involvement of an individual in an activity [25].

Regarding virtual words, immersion denotes the system’s ability to captivate the user, through both realism and interaction.

Realism is achieved mainly through visual and audio stimuli. It is not necessary for a virtual world to reproduce the real world in order for be immersive, but rather it has to be realistic and consistent. For example, 3D environments should not change the color palette very often, as an abrupt

change from bright colors to dark colors will break the user’s immersion. Consistency is also important for sounds. Immersion can be broken if the same object makes different sounds in similar contexts.

In a 3D environment, a user can walk and interact with the surrounding elements. This creates immersion by allowing free exploration. Depending on the amount and quality of the available interactions, the system can become more immersive. Multi-user systems allow a multitude of interactions and this is the reason they are generally more attractive.

Exploring a 360-degree photo or a video is less interactive, mostly because of the restricted movement and underlying technology and as a result, these systems are less immersive.

**Relationships between classification characteristics**

Type of technology	Availability	Immersion
360-degree photos and videos	High (Easily available online)	Low to Medium (Low number of possible interactions)
Single user 3D interactive environment	Medium to low (Can be available online)	Medium to High (Free movement and high number of interactions)
Multi-user 3D interactive environment	Low (Harder available online)	High (Free movement and very high number of interactions)

**Table 1. Availability and immersion based on type of technology used**

**EXAMPLES OF VIRTUAL WORLD SYSTEMS USED BY UNIVERSITIES**

Virtual world systems deployed by universities generally fall into one of three broad categories:

- 3D virtual tours;
- 3D campus simulations;
- 3D courses;

**Virtual 3D tours**

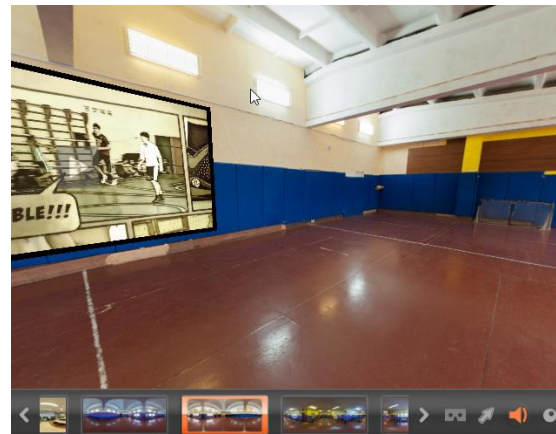
*Bonch-Bruevich Saint-Petersburg State University of Telecommunications Virtual Tour*

This virtual tour [5] uses spherical photos and videos to help students and prospective students to explore the university’s campus.

Most of the pictures are taken inside of the buildings and have a high resolution, enough to make small objects like electrical equipment easily distinguishable.

An interesting feature of this virtual tour is that it presents pictures taken in the laboratories and allows simple interactions with the objects inside them. For example, inside of the music laboratory, the user can click on instruments to make them play their specific sounds. This adds to a more interactive experience.

In addition, in some rooms, 2D videos that show students using them can be played, for example the fitness room. The videos look like being “projected” inside the room and rotate and scale accordingly as the user looks around the room.



**Figure 3. Saint-Petersburg University virtual tour interface with video being projected inside a 360-degree picture**

None of the taken 360-degree pictures captures human activity. This helps with capturing high-resolution still pictures, but makes the experience less immersive.

The tour is made using the Kolor [17] platform, which helps creating panoramas, virtual tours and 360-degree videos.

*Harvard University virtual tour*

The Harvard University virtual tour [6] uses 2D pictures of various buildings and landmarks as a main feature of the experience.

This tour presents the entire Harvard campus, so unlike the previously presented one, it does not feature pictures of the student laboratories. Most pictures show outdoor places, except some that are taken inside large buildings, such as the student hall or the stadium.



**Figure 4.** Harvard University virtual tour interface with “virtual tour guide” on the right side.

The experience is centered on locations and the user’s current location is displayed in a small map in the bottom left corner.

360-degree pictures and videos are also present, but are hidden away from the main interface. The user is notified when such a view exists nearby his current location and he can open it in an overlay above the interface.

An interesting feature of this tour is the “virtual tour guide”. This represents a cropped video of a Harvard student talking about the currently presented landmark, which adds to the user’s immersion. The tour guide’s narration is also displayed as scrolling text, which acts a subtitle.

As an augmentation to the online virtual tour, the users can download a virtual reality application to show the 360-degree pictures using a virtual reality helmet like Google Cardboard [18].

The application is powered by YouVisit.com [19], which helps creating virtual tours and 360-video presentations.

*Eindhoven University of Technology virtual tour*

The Eindhoven University of Technology virtual tour [7] is centered on 360-degree pictures of campus locations.

Unlike in the previous tours, in this one all the pictures feature human activity, as the application tries to immerse the user in the daily student life.

Some locations, for example parks, also feature sounds recorded near them, such as cars, birds and human voices, which also improves the user’s immersion.

A “virtual tour guide” is also present in some areas as a student in the 360-degree picture, and the user can click on a “play” button to hear what the student is saying.

Videos are used prominently throughout the tour, as the user can click on various objects representing university projects or activities and see a detailed video about them.



**Figure 5.** Eindhoven University of Technology virtual tour displaying a “virtual tour guide” (female student) integrated in the 360-degree picture. The “Play” displays a video about the presented vehicle.

Probably the most interesting part of this tour is a 360-degree video of the 2016 inauguration. The user can see in first person how the students are entering the hall, and the dean giving a presentation for them. Although the experience is not interactive, it is engrossing, especially since during the video a student “talks” to the user, creating a feeling of presence.

This tour is developed by the university without using a third-party platform.

**3D campus simulations**

*Virtual Campus of the Nanyang Technological University of Singapore*

The Nanyang Technological University of Singapore virtual campus [1] is the first of its kind. It was developed starting with 1997 by the team of Associate Professor Alexei Sourin.

It featured a full 3D simulation of the university campus and interactive lessons for 3D modelling.

The system was implemented using the Virtual Reality Modelling Markup Language (VRML) [20], a standard file format for representing 3D interactive vector graphics. Web browsers supported this format at the time, using it to display complex 3D models.

The entire campus model is still available online and it can be explored using specialized software for rendering VRML files.





**Figure 6. Nanyang Technological University of Singapore virtual campus. Snapshot of building interior.**

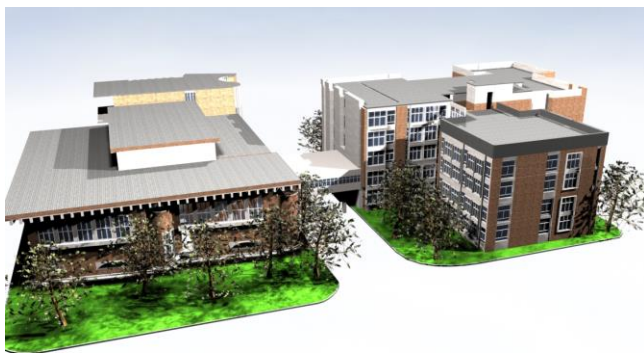
### 3D UPB

The “3D UPB” project is an ongoing virtual campus simulation project at the Politehnica University of Bucharest [32].

The project was started in 2010, as an initiative to create a complex educational and social platform. It was designed as a Massively Multiplayer Online Game (MMOG) and was intended to be a real time replica of the university’s campus. The simulator is meant to be extensible, with added components such as social-learning, mobile learning and serious games [33], [36]. Also, it should be integrated with the web based e-learning platform used by the university [34].

The initial architecture for the project was based on the open source Open Simulator [35] platform as an application server, with a MySQL database for storage.

Since the development of Open Simulator has slowed down, the team behind 3D UPB is looking for alternatives, for example Unity 5 [8].



**Figure 7. Render of the Faculty of Automatic Control and Computer Science buildings in 3D UPB**

### Virtual universities in Second Life

Second Life [21] is an online virtual world where users can socialize and interact with each other through their 3D

virtual avatars. This application belongs to the massively multiplayer online role-playing game (MMORPG) genre.

At its peak, in 2013, this virtual world had 1 million regular users, but since then this number fell to about 500 thousand users. The large number of players inspired universities to build virtual campuses in this game, using the application’s existing infrastructure to create “virtual social hubs” for their students [22].

The University of Hawaii built a virtual campus in 2010 [23]. The virtual space mimics the real university’s campus but also uses some other unrealistic elements such as a space rocket museum.

Students have used it as a social hub and as a place for 3D art exhibitions.



**Figure 7. University of Hawaii virtual campus in Second Life.**

Since Second Life as a platform lost much of its popularity, less students are currently using these virtual campuses but universities support them nonetheless.

### 3D courses

3D courses are widely available and cover a variety of subjects, like DNA simulation [8] and surgery [9].

Since some 3D courses can be used only in the universities laboratories, they can use the most advanced technology available.

A good example of this is the Bohemia Interactive Flight simulator, which puts the user in an actual cockpit, and uses a virtual reality helmet to render the 3D world and motion sensors to register the user input.



Figure 8. Student using a virtual reality flight simulator. [24]

### EMERGING TECHNOLOGIES

Presently there are two main emerging technologies in the context of virtual worlds:

- Virtual reality
- Augmented reality

#### Virtual reality

Virtual reality is a computer-generated simulation of a three-dimensional image or environment that can be interacted with in a seemingly real or physical way by a person using special electronic equipment, such as a helmet with a screen inside or gloves fitted with sensors [26].

Recently two types of virtual reality devices became available to a wide audience:

- Expensive, high fidelity devices like the Oculus Rift [27]. These devices have their own high-resolution screen and display vivid 3D environments, typically video games.
- Affordable, low fidelity devices like Google Cardboard [18] (less than 50 dollars). These devices use a smartphone screen as a monitor and provide the necessary optics to make a head mounted display. They are useful for short virtual reality experiences, like 360-degree videos.

With the increased popularity of such devices, more users can experience virtual reality than ever before.

Among the presented virtual world systems used by the universities, only the one from Harvard [6] uses a form of virtual reality. The 360-degree images presented in the tour are also available in a smartphone application that can be used together with a Google Cardboard device.

We consider that in the near future more virtual world systems will use virtual reality devices, as they provide a high degree of immersion.

#### Augmented reality

Augmented reality is a technology that superimposes a computer-generated image on a user's view of the real world, thus providing a composite view [28].

With this type of technology, developers can create parallel virtual worlds, for example video games where the users interact with real objects to see changes in the virtual environment.

Augmented reality is also used for educational purposes, as an informative overlay to the real world. A good example is the Star Chart [29] application, which shows a constellation map on top of a star system that the user is filming.



Figure 9. Star Chart application showing the Cygnus constellation [24]

Using augmented reality in university courses can increase student motivation [30], which is why we expect this type of courses to gain popularity.

### CONCLUSIONS

Universities around the world use virtual worlds as a tool for student engagement.

The systems presented in this paper are successful at achieving one or more of their purposes: presentational, social and educational.

We consider that the presented virtual tours have proven useful for engaging prospective students, as they provide an easy way to see the university campus and surroundings.

Virtual worlds have not been as successful at engaging active students, as the cost of creating an immersive 3D environment that supports multiple users has been very high until recently. With the appearance of frameworks such as Unity 5 [8] and Unreal Engine 4 [9] that allow much lower production costs, we consider that improved campus simulations will be available in the next years.

Emerging technologies that increase user immersion, for example virtual and augmented reality, can also be approached for generating student engagement.

## REFERENCES

1. Virtual Campus of the Nanyang Technological University in Singapore.  
*http://www3.ntu.edu.sg/home/assourin/vircampus.html* 1997
2. Virtual 3D Campus Design and Implementation. Hebei Polytech. Univ., Tangshan, China.  
*http://ieeexplore.ieee.org/document/5523195/*. 2010
3. A 3D Campus on the Internet – A Networked Mixed Reality Environment. National Taipei University.  
*https://link.springer.com/chapter/10.1007%2F978-3-642-03270-7\_19*. 2009
4. Design and Implementation of 3D Virtual Digital Campus - Based on Unity3D. Wuhan City Vocational College. *http://ieeexplore.ieee.org/document/7488531/*. 2016
5. Bonch-Bruevich Saint-Petersburg State University of Telecommunications Virtual Tour.  
*https://www.sut.ru/eng/3dtour*. 2017.
6. Harvard University virtual tour.  
*https://college.harvard.edu/admissions/visit/virtual-tour*. 2017.
7. Eindhoven University of Technology virtual tour.  
*https://virtualvisit.tue.nl/*. 2017.
8. Unity - Unity 5.0. [Internet]. Unity.  
*https://unity3d.com/unity/whats-new/unity-5.0*. 2017.
9. Unreal Engine – Unreal Engine 4. [Internet].  
*https://www.unrealengine.com/what-is-unreal-engine-4*. 2017
10. Simulation and visualization of a 3D biological environment. *https://www.ru.is/media/skjol-td/Mirco-Pazzaglia.pdf*. 2014
11. Virtual classroom courses creator. *http://edorble.com/*. 2017
12. Virtual 3D surgery courses.  
*https://www.touchsurgery.com*. 2017.
13. A History of Panoramic Image Creation [Internet]. Graphics.com. Available from:  
*http://www.graphics.com/article-old/history-panoramic-image-creation*. 2017.
14. WebGL - WebGL - OpenGL ES for the Web [Internet]. The Khronos Group. Available from:  
*https://www.khronos.org/api/webgl/*. 2017
15. WebAssembly [Internet]. Available from:  
*http://webassembly.org/*. 2017
16. File:Lone House.jpg. In: Wikipedia [Internet]. 2012 2017. Available from:  
*https://en.wikipedia.org/w/index.php?title=File:Lone\_House.jpg&oldid=470820077*
17. Kolor | Image-stitching, virtual tour, 360-degree video experts [Internet]. *http://www.kolor.com/*. 2017
18. Google Cardboard – Google VR [Internet].  
*https://vr.google.com/cardboard/*. 2017.
19. Home [Internet]. YouVisit. *https://www.youvisit.com/*. 2017.
20. VRML Virtual Reality Modeling Language [Internet].  
*https://www.w3.org/MarkUp/VRML/*. 2017
21. Official Site | Second Life [Internet].  
*http://secondlife.com/*. 2017.
22. Second Life Education Directory - Second Life Wiki [Internet]. Available from:  
*http://wiki.secondlife.com/wiki/Second\_Life\_Education\_Directory*. 2017.
23. Second Life [Internet]. Available from:  
*http://www.hawaii.edu/secondlife/*. 2017.
24. Bohemia Interactive Simulations [Internet].  
*https://bisimulations.com*. 2017.
25. Immerse Meaning in the Cambridge English Dictionary [Internet].  
*http://dictionary.cambridge.org/dictionary/english/immerse*. 2017.
26. Virtual reality - definition of virtual reality in English [Internet].  
*https://en.oxforddictionaries.com/definition/virtual\_reality*. 2017.
27. Oculus Rift | Oculus [Internet].  
*https://www.oculus.com/rift/*. 2017.
28. Augmented reality - definition of augmented reality in English. [Internet]  
*https://en.oxforddictionaries.com/definition/augmented\_reality*. 2017.
29. Star Chart. [Internet]  
*https://www.engadget.com/2013/08/18/star-chart-is-a-nifty-and-free-augmented-reality-astronomy-app/*. 2017.
30. Di Serio Á, Ibáñez MB, Kloos CD. Impact of an augmented reality system on students' motivation for a visual art course. *Computers & Education*. 2013
31. Google Maps [Internet] *https://maps.google.com*. 2017
32. 3D UPB - Politehnica University of Buchares. [Internet] *http://3d.upb.ro/*. 2017
33. Moldoveanu A, Gradinaru A, Ferche O-M, Stefan L. The 3D UPB mixed reality campus: Challenges of mixing the real and the virtual. In: System Theory, Control and Computing (ICSTCC), 2014 18th International Conference [Internet]. IEEE; 2014. p. 538–543. Available

from:

<http://ieeexplore.ieee.org/abstract/document/6982472/>

34. Moldoveanu AD, Moldoveanu F, Asavei V, Egner A, Morar A. From HTML to 3DMMO–A Roadmap Full of Challenges. *Advances in Intelligent Control Systems and Computer Science*. 2013;379–392.

35. OpenSimulator [Internet].

[http://opensimulator.org/wiki/Main\\_Page](http://opensimulator.org/wiki/Main_Page). 2017

36. Ștefan L, Moldoveanu F, Moldoveanu A. BLENDED LEARNING IN A MIXED REALITY-BASED 3D MULTI-USER VIRTUAL ENVIRONMENT. *Conference proceedings of »eLearning and Software for Education« (eLSE)*. 2014;(1):105–12.

Link to word version:

[http://slbucur.me/rochi2017/Overview\\_FINAL.doc](http://slbucur.me/rochi2017/Overview_FINAL.doc)

# Agile Software Development Process with Task Models having Triggers, Preconditions and Context Constraints

Peter Forbrig

University of Rostock  
Albert-Einstein-Str. 22,  
18051 Rostock  
peter.forbrig@uni-rostock.de

Gregor Buchholz

University of Rostock  
Albert-Einstein-Str. 22,  
18051 Rostock  
gregor.buchholz@uni-rostock.de

## ABSTRACT

Agile software development asks for seamless integration of software engineering methods and approaches from HCI. A process model will be discussed that includes HCD as one sprint ahead in the corresponding process model. Additionally, model simulations are suggested for improved evaluation of design decisions.

The language CoTaSL (Cooperative Task Specification Language) is introduced for specifying the activities of users and certain context aspects. Its editor was developed based on the Xtext framework. The language is a human-centered form of CoTaL (Cooperative Task Language) that can be simulated by CoTaSE (Cooperative Task Simulation Environment). CoTaL and CoTaSE will be discussed as well.

Different approaches for specifying collaborative work are presented. Advantages and disadvantages are discussed. An outlook is provided that suggests some further support for CoTaSL by other tools.

## Author Keywords

Agile Development Process, Task model, Constraints, Context, Coordination model, Team model, Role model.

## ACM Classification Keywords

H.5 [Information Interfaces and Presentation]: H.5.2 User Interfaces – User-centred design; H.5.3 Group and Organization Interfaces – Theory and models.

## General Terms

Human Factors; Requirements Specification, Simulation

## INTRODUCTION

Currently, it is common ground that software should be developed iteratively in an agile way. This is necessary to improve the communication with the stakeholders. Otherwise, it is difficult or even impossible to capture the correct requirements. However, the classical waterfall model has some advantages for managing milestones. Therefore, a combination of both approaches makes sense that supports both the management aspect and the software development aspect.

In this paper, we will discuss the usage of extended task models that together with corresponding tool support can support a human-centered agile development process. The

language CoTaSL was developed to allow specifications of collaborative activities. This is especially important for developing smart collaborative environments and workflow applications.

Each CoTaSL specification consists of one team model and several role models. The simulated team model instance reflects the state of the collaboration of the different role model instances.

First, we will focus on the integration of HCI aspects into a process model of agile software development. After that, the language CoTaSL will be introduced based on CoTaL and different simulation situations in CoTaSE.

## SOFTWARE DEVELOPMENT PROCESS MODELS

Agile software development has become popular since the 1990s because many projects failed. The reason for that was that it took too much time from finalized requirements specification to first tests of the developed system. At the beginning of the 2000s a manifesto [1] was published that characterizes the agile idea by twelve main principles.

An overview of influences on agile software development is presented by Fig. 1. It shows approaches from planning, analysis, design, build, test, deploy, and review that found their way into the agile approach.

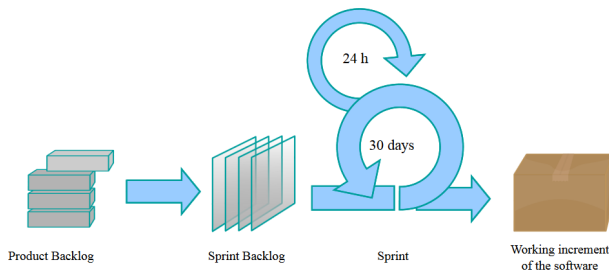


Figure 1. Influences to agile software development<sup>1</sup>

<sup>1</sup><http://blog.kiandra.com.au/wp-content/uploads/2013/08/Agile-21.png>



Currently, one of the most popular agile software development approaches is SCRUM [19]. It is quite good characterized by the following Fig. 2.



**Figure 2. Model for the SCRUM development process**

Originally, the process model of SCRUM (like software engineering in general) focused on software development activities only. However, usability aspects are important as well. Therefore, the integration of HCI aspects into the development process is discussed by several authors.

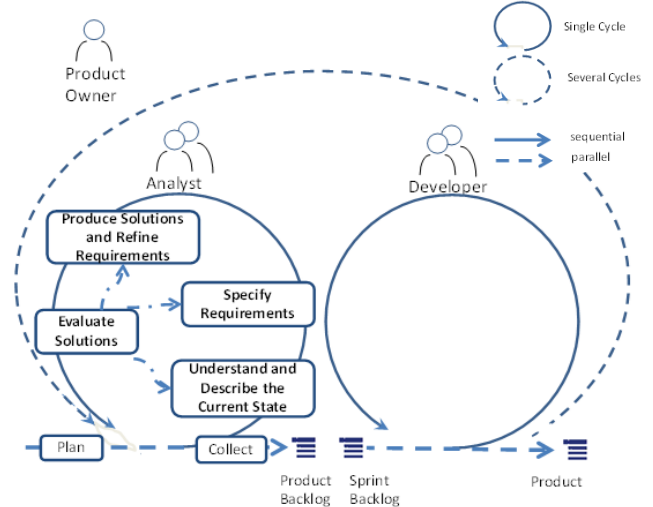
Two interleaving processes for developers and HCI specialists are suggested by Sy [21]. She suggest the gathering of some user data and the establishment of a common plan at the very beginning of a project. This has to be done by all participants. Developers start in the first development cycle with implementations that are independent from the user interface. This could e.g. be certain internal services of the application. At the same time, HCI specialists provide design solutions for cycle two and gather customer data for cycle three.

Developers implement the design solutions from cycle one in cycle two and in parallel their code from cycle one is tested by HCI experts. Additionally, they design for the next cycle and analyze for the cycle after the next cycle. This is the general development pattern. In this way, HCI specialists work two cycles ahead to developers in analyzing customer data and one cycle ahead in developing design solutions. Paelke et al. [16] published another process model for combining agile software development and agile user-centered activities.

HCD (Human-Centered Design) [12] is popular among usability and user experience experts in the same way as SCRUM with software developers. It provides an iterative approach for requirements gathering and considers the context of use, the requirements of the users and the evaluation of design solutions. The HCD process has been standardized by ISO 9241-210. Fig. 3 presents a part of an agile process model from [9] that includes an adapted HCD process model. It combines the ideas of HCD, Sy [21] and Paelke et al. [16]. A first version of the process model was presented in [8].

Currently user stories, scenarios, or use-case slices are used to capture requirements of applications. However, sometimes more elaborate activities for specifying activities

are helpful to explore an application domain. Task models have been proven a useful tool for HCD. However, sometimes detailed features of the models are missing.



**Figure 3. HCD details for SCRUM model sprints**

The next paragraph will discuss a special version of task models that allows the creation of instances and the specification of context dependencies. This is especially important for specifications in smart environments or applications following a complex workflow.

### SPECIFICATION OF SEQUENCES OF ACTIONS

#### Task Model

Sequences of actions can be specified by use cases [22] or business process models in the notation of BPMN [4] or S-BPM [6]. However, in the field of software ergonomics task models have proved to be helpful specifications [17]. TKS [13] and HTA [2] were one of the first approaches to specify possible task executions by models. Currently, systems like CTTE [5] and HAMSTERS [11] exist that allow the generation of other models and the cooperation with different other systems. Additionally, they allow the simulation of task models. Collaboration specification was discussed in [14], [18] and [20]. This allows very good stakeholder involvement at early stages of software development, which can be considered to be a substantial contribution to user centered development methodologies.

However, our application domain of “Smart Environments” asks for features that neither the cooperation model of CTTE nor the task groups in HAMSTERSs fulfill. This includes, first and foremost, the opportunity to have a variable number of task model instances. It is not known before the actual execution how many actors will collaborate. The current systems do not have this flexibility. Additionally, we require the temporal relation “instance iteration” that means the second iteration of a task can be started before the first ends. Task models need this feature for practical applications. The importance of models was already discussed in [7].



The specification language CoTaL (Cooperative Task Language) and a corresponding interpretation environment CoTaSE (Cooperative Task Simulation Environment) were developed for this reason. Some ideas were taken from the predecessor language CTML [23]. The separation of task models into several role models and one team model is one of the ideas of CTML that we included into CoTaL. The activity of a certain kind of actor (in terms of use cases), subject (in terms of S-BPMN) or pool (in terms of BPMN) is modelled as a role. During runtime several instances of a role model can exist. This is of course also true for a simulation. The instances specify the state of the task performance of a specific person (actor instance). The instantiation of a task model immediately delivers a stateful “running” model instance. This is very similar to the relation of classes and objects in object-oriented programs.

The cooperation of related role instances is reflected by the instance of a team model. This model may vary in its level of detail. It is thus possible to reflect each and every task of all role models. However, it is also viable to have one single team task only. The state of this task model shows when the collaborative activities are running and when they are finished. The optimal solution seems to be something in between those two extremes.

The two kinds of task models will be discussed within the following paragraph by an example, which assumes that a conference session has to be organized in a smart meeting room. It is further assumed that there are only the two roles of “Chair” and “Presenter”. The chair has to introduce a session first, followed by the presenters. After all presentations of the conference session are finished a joined discussion will be performed. We will first have a look at the structure of the team model and will later discuss the structure of role models.

### Team Model

The cooperation of actors is specified by a team model. The model specifies all roles involved and specifies events coming from individual task model instances. Activities of actors can be represented by different task model instances simultaneously. For example, a person can have the role of a chair and the role of a presenter at the same time.

Let us have a look at a simple example of performing a conference session. A chair has to manage the presentations first and afterwards the discussion. The task “Manage presentation” consists of two subtasks “Introduce talk” by the chair followed by “Present” assigned to the presenter.

In the following two visual representations of the corresponding team will be shown. They do not show who executed or is intended to perform a tasks. The visualization of task tree instances uses different colors and symbols to illustrate the state of the respective task. The set of possible task states include *disabled* (✘, red), *enabled* (○, yellow), *running* (8, green), *skipped* (↯, gray) and *finished* (✓, white). Please keep in mind that only leaf tasks can be

started directly. Therefore, inner nodes are depicted in a slightly paler colour shade. Their state is derived bottom-up from their respective child tasks. Fig. 4 provides a visualization of a model instance that consists of the team model.

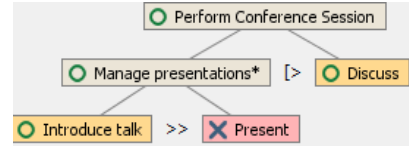


Figure 4. Session example of a task tree instance.

According to the model instance’s current state, a talk can be introduced or the discussion can be started. A talk has to be introduced first and can only be presented afterwards (>>, enable). To perform a conference session several presentations have to be managed (\*, iteration). The sequence of presentations is stopped ([>, deactivate) by starting the discussion. Fig. 4 presents the task visualization like CTTE. We will call it task tree. An alternative presentation like in HAMSTERS, where temporal operators are represented as nodes is shown in Fig. 5.

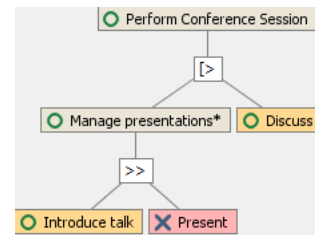


Figure 5. Session example of a syntactic tree instance.

The visualization of the task model in Fig. 5 is called syntactic tree. Within the CoTaSE tool both representations of task model instances can be shown alternatively or even together. The XML specification of the team model is the following one.

```
<?xml version="1.0" encoding="UTF-8"?>
<taskmodel name="Performing a conference" role="Team">
  <roles>
    <role name="Chair" file="Chair.xml"/>
    <role name="Presenter" file="Presenter.xml"/>
  </roles>
  <task name="Perform Conference Session">
    <task name="Manage presentations"
      operator="disabling" iterative="true">
      <task name="Introduce talk" operator="enabling"
        startTrigger=
          "Chair.oneInstance.IntroducePresenter.start"
        endTrigger=
          "Chair.oneInstance.IntroducePresenter.end">
      </task>
      <task name="Present"
        startTrigger=
          "Presenter.oneInstance.IntroduceTalk.start"
        endTrigger=
          "Presenter.oneInstance.EndTalk.end">
      </task>
    </task>
    <task name="Discuss"
      startTrigger="Presenter.allInstances.EndTalk.end"
      endTrigger="Chair.allInstances.CloseSession.end">
    </task>
  </taskmodel>
```

For each leaf task in the team model there exists a start trigger and an end trigger. Triggers are specified in an OCL-like notation [15]. The first part of the expression is a role name that is followed by “oneInstance” or “allInstances”. This is followed by a task that is started or ended. The task “Introduce talk” is started when an instance of the role “Chair” has started the task “Introduce presenter” (*Introduce presenter* is equal to *IntroducePresenter*). It ends when the introduction is ended. A presentation is started, when one presenter started to present his or her talk. The discussion starts when all talks were presented. The conference session ends when all “Chairs” closed the session.

To avoid the XML details of CoTaL the domain specific language CoTaSL (Cooperative Task Specification Language) was designed based on Xtext [25]. Xtext generates a syntax driven text editor that can be combined with Xtend [24] to generate CoTaL specifications.

The above example of CoTaL looks like the following specification.

```
team for t {
  task Perform_Conference_Session = Manage_presentation{*} [> Discuss;
  task Manage_presentation = Introduce_talk >> Present;
  task Introduce_talk;
  trig Introduce_talk = ( start.Chair.oneInstance.Introduce_presenter.start );
  trig Introduce_talk = ( end.Chair.oneInstance.Introduce_presenter.end );
  task Present;
  trig Present = ( start.Presenter.oneInstance.Introduce_talk.start );
  trig Present = ( end.Presenter.oneInstance.Give_talk.end );
  task Discuss;
  trig Discuss = ( start.Presenter.allInstances.Give_talk.end );
  trig Discuss = ( end.Chair.oneInstance.Close_session.end );
}
```

It is assumed that it is much easier to use the domain specific language CoTaSL than CoTaL, which is optimized for machine interpretation while CoTaSL is intended for human usage.

### Role Model

A role model specifies the activities of one role. In our example, we need one for the role “Chair” and one for the role “Presenter”. The XML specification of the task model of a “Chair” will be discussed first.

```
<?xml version="1.0" encoding="UTF-8"?>
<taskmodel name="Chairing a session" role="Chair">
  <task name="Chair session">
    <task name="Introduce session"
      operator="enabling">
    </task>
    <task name="Introduce presenter"
      operator="disabling"
      iterative="true">
    </task>
    <task name="Open discussion"
      operator="enabling"
      precondition="Presenter.allInstances.EndTalk">
    </task>
    <task name="Close session"/>
  </task>
</taskmodel>
```

The above specification uses a precondition that has a similar notation to that of triggers. In preconditions, the completed execution of the mentioned task is assumed.

Therefore, the information “.start” or “.end” at the end of the expression is missing.

According to the model, a chair introduces a session and several presenters afterwards. After all presentations, the discussion is opened and finally the chair closes a session. A precondition for opening the discussion is the fact that all presenters have finished their talks. Fig. 6 visualizes the corresponding task hierarchy and presents the start state of the task instance.

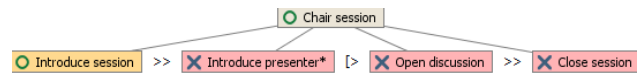


Figure 6. Task model instance of a chair.

In CoTaSL these activities can be specified as follows:

```
role Chair for Chair {
  task Chair = Introduce_session >> Introduce_presenter{*} >> Open_discussion >>
    Close_session;
  pre Close_session -> (Presenter.allInstances.End_talk );
}
```

Let us assume that a presenter introduces his or her talk first and afterwards, a series of slides is explained before the talk is ended. A slide is explained by first activating it, taking an optional nip of water afterwards and finally explain the slide. Fig. 7 provides a corresponding animated task model instance.

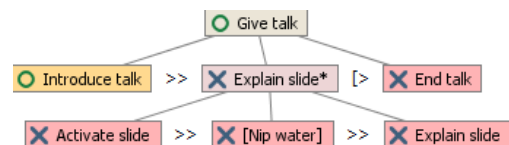


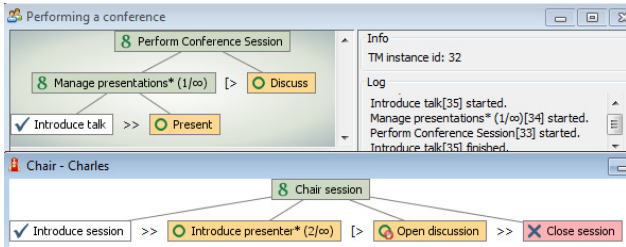
Figure 7. Task model instance of a presenter after instantiation.

Below you will find the CoTaL specification of the task model of the role “Presenter”. A talk can be introduced after a chair introduced a new presenter.

```
<?xml version="1.0" encoding="UTF-8"?>
<taskmodel name="Presenting a talk" role="Presenter">
  <task name="Give talk">
    <task name="Introduce talk"
      operator="enabling"
      precondition="Chair.oneInstance.IntroducePresenter">
    </task>
    <task name="Explain slide"
      operator="disabling" iterative="true">
      <task name="Activate slide" operator="enabling"/>
      <task name="Nip water" operator="enabling"
        optional="true">
      </task>
      <task name="Explain slide"/>
    </task>
    <task name="End talk"/>
  </task>
</taskmodel>
```

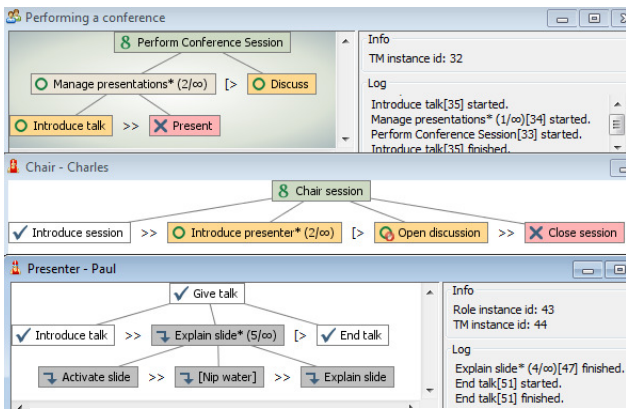
There seems to be no further discussions of the role model of a presenter necessary. No new language features were used.

In the following example we will have three presenters Paul, Peter, and Penny and two chairs Charles and Chris. The following Fig. 8 presents the visualization of the state of the simulated team and role model instances of chair Charles.



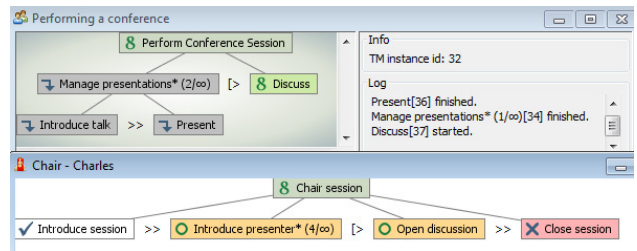
**Figure 8. Model instances after scenario “Introduce session” and “Introduce presenter”.**

One can see the thrown events in the log information of each model instance. In Fig. 8 only the team model logs are presented because of lack of space. For Charles the task “introduce presenter” is already executed once and now the second activation waits. From the team model instance one can see that yet no presentation was given. However, a presentation could start right now. After Paul performed his talk with four slides (the fifth could have been presented), the state of the model instances presented in Fig. 9 is reached. The team model instance and the role model instance of Charles reflect the fact that the second talk can be introduced.



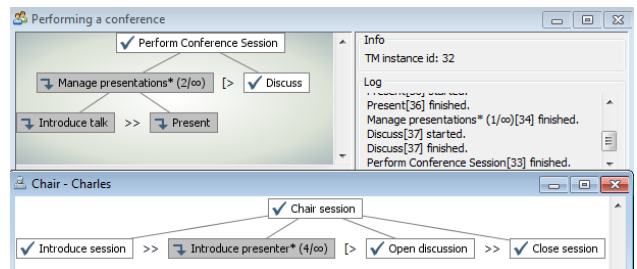
**Figure 9. Model instances after the first talk**

According to the temporal relations of the task model chair Charles can open the discussion. However, there is a constraint that is still not fulfilled. The constraint is the precondition “Presenter.allInstances.endTalk”, which means that the discussion cannot be started until all presenters finished their talks. The red overlay icon in Charles’s task “Open discussion” symbolizes this. We now assume that all presenters finished their talks and chair Charles can open the discussion. This situation is reflected by Fig. 10. The task models of the presenters are omitted because they all look like that of presenter Paul in Fig. 9.



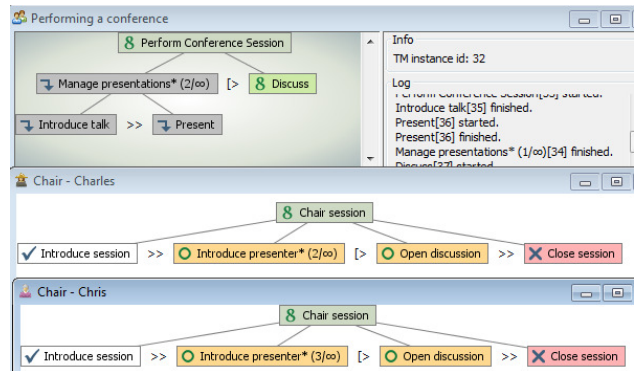
**Figure 10. Model instances after all talks.**

Charles can open the discussion at this stage. No new presenters can be introduced because all presenters were active already. “Introduce presenter” is in the fourth iteration. Fig. 11 presents the final states for the model instance of the conference and the chair.



**Figure 11. Final model instances**

The scenario of the conference session went well. However, the specified constraints were very general. One instance of a presenter had to start a talk and one instance had to end it. Indeed, it is not checked whether this is the same instance. The following situation would be possible.



**Figure 12. Model instances after three talks**

In the example of Fig. 12 Charles and Chris introduced a session. Additionally, Charles introduced one presenter and Chris introduced two. Both can open the discussion now. However, this is not the way a session should be performed. The chair who opened a session should also close it. Additionally, there should not be two openings. To allow a more precise specification the concept of context and related binding of variables is used.

## Context and Binding Variables

Sometimes it is necessary to specify certain constraints. A session should e.g. be closed by those chair that opened it. The constraint could be specified by OCL [15] expressions. However, their readability is limited. Based on the experience with CTML [23], a very reduced version of OCL like expressions is used. Let us have a look at the extended example of the team model.

```

team for t {
  contextVar Session { roleVar (Chair) C; };
  contextVar Talk { roleVar (Presenter) P; };
  task Perform_Conference_Session = Introduce_session >>
    Manage_presentation{*} [> Discuss;
  contextBind Perform_Conference_Session -> Session;
  task Introduce_session;
  roleVarBind introduce_session -> C.introduce_session;
  task Manage_presentation = Introduce_talk >> Present;
  contextBind manage_presentations -> Talk;
  task Introduce_talk;
  roleVarBind Introduce_talk -> C.Introduce_presenter;
  task Present;
  roleVarBind present_talk -> P.give_talk;
  task Discuss;
  roleVarBind open_discussion -> C.open_discussion;
}

```

Contexts can be defined for certain tasks in the team model:

```

contextBind manage_conference -> Session;
contextBind manage_presentations -> Talk;

```

Linked to such contexts it is possible to introduce variables for roles:

```

contextVar Session { roleVar (Chair) C; };
contextVar Talk { roleVar (Presenter) P; };

```

The value of such a variable can be bound. The binding of values within such a context takes place by a certain task execution of an instance of the specified role:

```

roleVarBind introduce_session->C.introduce_session;

```

This would mean that the instance of any role model that executes first the task "Introduce Session" is stored into the variable "C". Later use of this variable will refer to the stored instance.

By specifying contexts, activities can be bound to a certain subject (instance of a role). In the above example, the chair that introduces a session, introduces the presenter and closes the session has always to be the same person (role instance). Additionally, within the example specification presenters are bound to a talk. However, there is a need for more dynamic bindings. Otherwise, the first presenter had to give all talks. Within the iteration every time the presenter that first started his talk is bound to the corresponding variable. This is possible by a specific attribute (`bindAllIterations->"false"`). The default value of this attribute is "true" and means that the first bound values are the same for all iterations.

With this improved specification with an additional team task ("Open session") and the specified constraints, the situation presented in Fig. 10 looks now like shown in Fig. 13. The chair Charles cannot introduce a new presenter after one presenter was introduced. In the old version of the

specification, the chair was able to introduce the next talks while a presenter was still showing slides. This is no longer possible with the new extended specification. For both versions of the specification, it is possible within the CoTaSE environment to create new instances of role models. In this way, a new presenter can be created. This reflects the fact that a new subject enters the room that has the corresponding role. This instance creation, of course, influences the states of the model of chair Charles and the model of the team.

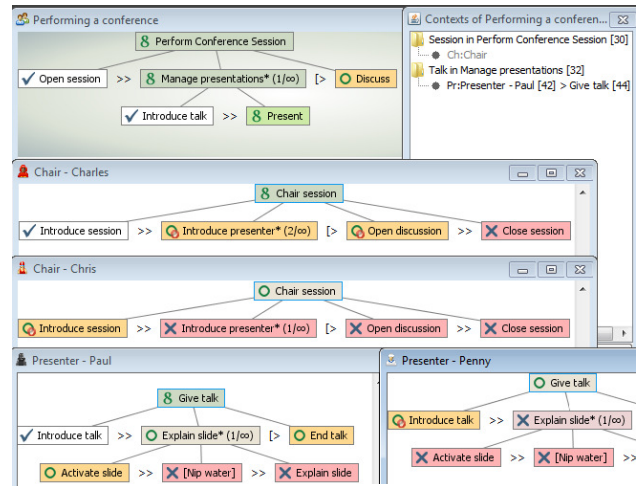


Figure 13. State after start of the first presentation with context binding

In the upper right corner, one can see that Paul is the current presenter that is stored in `Pr`. Charles is the chair that started the session. Neither Charles nor Chris are at this stage allowed to introduce a new talk. Chris has to wait until the whole session is over and Charles has to wait until the current presentation is finished.

However, in certain circumstances it makes sense to distinguish between iteration and instance iteration. Instance iteration allows the beginning of the next iteration before the preceding iteration ended.

The next paragraph will present some details of instance creation and instance iteration.

## Dynamic Instances and Instance Iteration

The environment CoTaSE allows the creation of task model instances during runtime, in our case simulation time. Fig. 14 shows the simulation state after the presentations of Paul and Penny (all presenters). Fig. 15 presents the states of the model instances after creating the presenter Peter. Chair Chris cannot open the discussion anymore because there is the constraint that all presenters ended their talks. However, he can introduce a presenter. This was not possible without the new presenter. The team model changes as well. The management of presentations becomes active again and "Introduce talks" is not finished anymore.



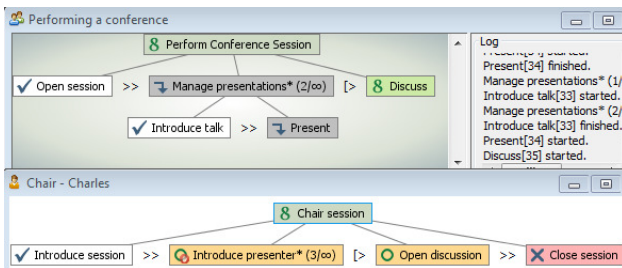


Figure 14. Situation after all two presentations

Experimenting with task models resulted in some further needs related to instances. From our point of view it is necessary to provide support for instance iterations. Sometimes it makes sense to start with the next iteration before the first one is finished. For the domain of conference session it might be possible that a chair manages several online presentations in parallel. In this case the second presentation can start before the first one finishes. It is assumed that at most four presentations can run in parallel. Fig. 16 presents the situation after the first presenter was introduced and started and the second one was introduced. Instance iterations (marked with a “#” followed by up to two integers specifying optional minimum and maximum instance numbers) are used in the team model and in the model of the chair. Instances of the tasks within the iterations are executed in parallel. The interleaving operator (|||) is used to visualize this fact.

introduce the third presentation. While Paula already introduced her talk and explains slides, Jan is able to introduce his talk. The third presenter Peter is not included in Fig. 16 because his state is identical to that of Fig. 15. He cannot introduce his talk because the chair did not introduce the third talk of the conference session.

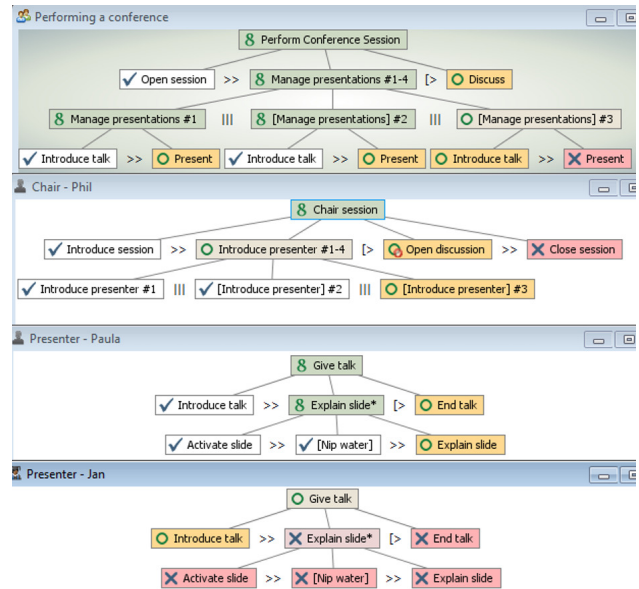


Figure 16. Task model instances with instance iteration.

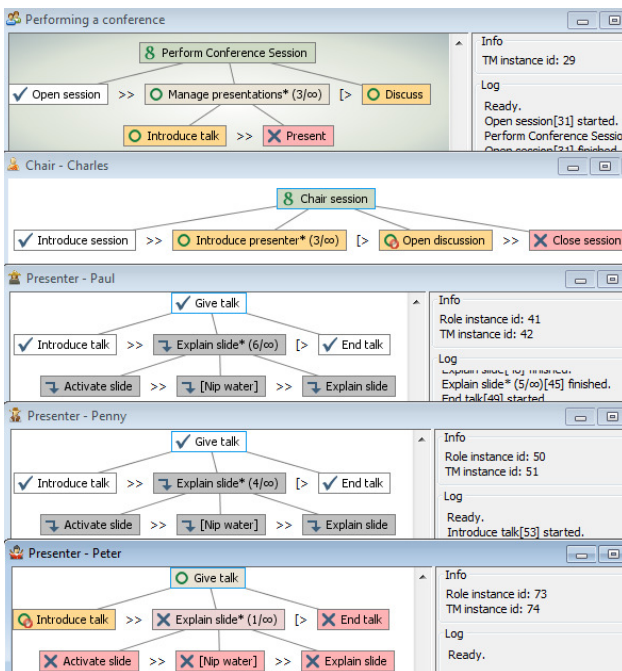


Figure 15. Situation after creating presenter Peter in the state provided in Fig. 14.

One can see in the team model of Fig. 16 that two presentations were introduced. The same fact can be concluded for the model of the chair Phil. He is able to

## SUMMARY

It was suggested in this paper that models for describing activities of users help to integrate HCD into agile software development. A process model was presented that suggests HCD activities in one sprint ahead to the development.

The language CoTaSL was introduced as a domain specific language for task models. The language allows very precise specifications, allows context bindings and instance iterations. Additionally, CoTaSL allows abstract and human-centered specifications.

In [10] one can see how the principles of such a language (it was CoTaL there) allow subject-oriented business process specifications. A kind of workflow system based on this concept is presented in [3]. It allows distributed collaborative simulation of activities over the Internet.

Currently, tool support is provided for CoTaSL by a syntax-driven editor and a generator to CoTaL. The generated code can be simulated within the environment CoTaSE. However, the language can be transformed into a CTT specification or the notation of HAMSTERS.

For the future, the inclusion of more features into the language might be useful. In this way, it can become a general programming language for task models. It might even be possible to find another even more abstract domain specific language that can be translated to CoTaSL.

## ACKNOWLEDGMENTS

The work was supported by DFG graduate school 1424 Multimodal Smart Appliance Ensembles for Mobile Applications (MuSAMA) at the University of Rostock, Germany.

## REFERENCES

1. Agile manifest: <http://agilemanifesto.org/iso/en/principles.html>, last visited, July 5<sup>th</sup> 2017.
2. Anett, J.: Hierarchical Task Analysis. In Diaper D., Stanton N. (Eds), *The Handbook of Task Analysis for Human-Computer Interaction* (pp. 67-82). Lawrence Erlbaum Associates, 2004.
3. Buchholz, G., and Forbrig, P.: Extended Features of Task Models for Specifying Cooperative Activities, accepted for EICS 2017, Lisbon.
4. BPMN: <http://www.bpmn.org/> last visited, July 5<sup>th</sup> 2017.
5. CTTE: <http://giove.cnuce.cnr.it/ctte.html>, last visited January 6<sup>th</sup> 2017.
6. Fleischmann, A., Schmidt, W., Stary, C., Obermeier, S., and Börger, E., 2012: *Subject-Oriented Business Process Management*. Springer, ISBN 978-3-642-32391-1, pp. I-XV, 1-375
7. Forbrig, P.: Interactions in Smart Environments and the Importance of Modelling, *Proceedings of the National Conference on Human-Computer-Interaction - Romanian Journal of HCI*, 2012, <http://rochi.utcluj.ro/rrioc/articole/RoCHI-2012/RoCHI-2012-Forbrig.pdf>
8. Forbrig, P., and Herczeg, M.: Managing the Agile Process of Human-Centred Design and Software Development, *IFIP WG 13.2 Workshop proceedings: User Experience and User-Centered Development Processes*. INTERACT 2015. Bamberg.
9. Forbrig, P.: When Do Projects End? - The Role of Continuous Software Engineering. *Proc. BIR 2016*: 107-121
10. Forbrig, P., and Buchholz, G.: Subject-Oriented Specification of Smart Environments, *Proc. 9th Conference on Subject-oriented Business Process Management, S-BPM ONE 2017*, Darmstadt, Germany, March 30-31, 2017. ACM 2017, ISBN 978-1-4503-4862-1, <http://dl.acm.org/citation.cfm?id=3040570>
11. HAMSTERS: <https://www.irit.fr/recherches/ICS/software/hamsters/>, last visited January 6<sup>th</sup> 2017.
12. HCD: <https://www.iso.org/standard/52075.html>, last visited May 20<sup>th</sup>, 2017
13. Johnson, P., Johnson, H., Waddington, R., and Shouls, A.: *Task-Related Knowledge Structures: Analysis, Modelling and Application*. BCS HCI 1988: 35-62.
14. Martinie, C., Barboni, E., Navarre, D., and Palanque, P. A., Fahssi, R., Poupart, E., and Cubero-Castan, E.: Multi-models-based engineering of collaborative systems: application to collision avoidance operations for spacecraft. *EICS 2014*: 85-94
15. OCL: <http://www.omg.org/spec/OCL/>, last visited January 6<sup>th</sup> 2017.
16. Paelke, V. and Nebe, K. Integrating Agile Methods for Mixed Reality Design Space Exploration. In *Proceedings of the 7th ACM conference on Designing interactive systems (DIS '08)*. ACM, New York, NY, USA, 240-249. <http://doi.acm.org/10.1145/1394445.1394471>
17. Paternò, F.: *Model-Based Design and Evaluation of Interactive Application*, Springer Verlag, ISBN 1-85233-155-0.
18. Penichet, V. M. R., Lozano, M. D., Gallud, J. A., and Tesoriero, R.: *Task Modelling for Collaborative Systems*. TAMODIA 2007: 287-292
19. SCRUM: [https://en.wikipedia.org/wiki/Scrum\\_\(software\\_development\)](https://en.wikipedia.org/wiki/Scrum_(software_development)), last visited May 20<sup>th</sup> 2017.
20. Solano, A., Toni Granollers, T., Collazos, C. A., and Rusu, C.: Proposing Formal Notation for Modeling Collaborative Processes Extending HAMSTERS Notation. *WorldCIST (1) 2014*: 257-266
21. Sy, D.: Adapting usability investigations for agile user-centered design. *J. Usability Stud.* 2 (3), pp. 112-132, 2007.
22. UML: <http://www.uml.org/>, last visited January 6<sup>th</sup> 2017.
23. Wurdel, M., Sinnig, D. and Forbrig, P. 2008. CTML: Domain and Task Modeling for Collaborative Environments. In: *Journal of Universal Computer Science 14(19) (Special Issue on Human-Computer Interaction)*, 3188-3201.
24. Xtend: <http://www.eclipse.org/xtend/>, last visited May 20<sup>th</sup>, 2017.
25. Xtext: <https://eclipse.org/Xtext/documentation/>, last visited May 20<sup>th</sup>, 2017.



# Perspective-Based Usability Inspection for Active Aging in Place Websites

Adriana Alexandru, Marilena Ianculescu, Maria Gheorghe-Moisii

National Institute for Research and Development in Informatics - ICI Bucharest

Bd. Maresal Averescu, nr.8-10, Bucharest, Romania  
{manina, adriana, moise}@ici.ro

## ABSTRACT

The effects of the ongoing and increasing aging population on the social and economic life have imposed a shift towards creating new ways to adapt the society. Elderly have become a group of people with enhanced demands and necessities. In order to release the financial burden on the health and social systems, an age-friendly approach of designing and implementing better adapted digital technology is a sustainable support for an active and independent life in the personal residence, namely aging in place. The user interface is a key element for making a web application suitable and accepted by elderly. Perspective-based technique for usability inspection is particularly beneficially for identifying these types of problems by ability of diverse typology of evaluators. This paper presents two case studies of websites for aging in place; the websites' usability is analysed with the above technique. The envisioned further improvement and development based on Big Data Analytics of the two websites should be done taking into consideration the recommendations based on the highlighted age-related general usability problems.

## Author Keywords

User interface; elderly; Big Data; aging in place; usability

## ACM Classification Keywords

H5.2; H-3.5; J3.

## INTRODUCTION

The world's population is getting older due to the lower birth rates and longer and healthier lives. The aging population is the biggest consumer of the resources of longterm care, health, and social systems. Avoiding premature and unnecessary institutionalisation can relieve the financial burden and put forward a better and dignified life of an elderly and his/her family.

*Aging in place* has been defined as "the ability to live in one's own home and community safely, independently, and comfortably, regardless of age, income, or ability level" [4].

Aging in place implies the use of a large variety of smart devices such as embedded in-home and wearable sensors, assistive technology able to sustain the preservation of an active life on longterm and web applications aiming to support the management of chronic health conditions.

Due to the degenerative diseases, elderly undergo an increased number of challenges when they use a web

application. Furthermore, even among the particular area of the elderly group, users have distinctive degrees of analytical competence, health and information and communications technology (ICT) literacy; therefore they can process and interpolate the same health information in a different manner. The user interface should be age-friendly to suit larger categories of users with degenerative problems or ICT skills.

*Usability testing* plays an essential role in implementing web applications that target the aging population. Identifying the issues that might restrain the acceptance and use of digital technologies by the elderly (even from early stages of software development) has demonstrated to be efficient and able to maximize the necessary adaptations to the older people's demands.

There late guidelines specifically applied to the design of web applications for elderly actively adjust and enlarge the user's accessibility and know-how [14].

Among the usability testing methods, the *perspective-based inspection* can provide a better approach for an elderly-centered evaluation of the web applications.

Big Data has rapidly imposed in healthcare due to the way it has revolutionized the exploitation of the massive stores of data and knowledge. That proved to be of a great importance, since in 2012 the Poneman Institute estimated that 30% of all the electronic data storage in the world was occupied by the healthcare industry [12]. Regarding the aging in place, Big Data has the potential to sustain it by an enlarged healthcare infrastructure that supports the healthcare specialists and providers to identify, handle, analyze, and exploit huge amounts of data in order to provide a better personalized degenerative diseases management.

This paper aims to use the results of usability inspections of two websites that address issues related to aging in place domain and which are presented in two *case studies*. Different types of inspectors have been chosen as the source of different perspectives, including the ones of the elderly. The results of this approach led to the statement of some recommendations for further development of the web applications and for a better design of new eHealth solutions targeting the aging in place.

## **AGING POPULATION AND DESIGNING APPROPRIATE USER INTERFACES**

The personal residence is an environment where the elderly should feel safe, active, efficient and independent. Adapting and shaping the web applications into something familiar should be a well-defined goal in designing the interfaces. Their design should be based on affective, aged and home associated values and on appropriated standards like ISO/IEC Guide 71 that “has successfully addressed the importance of being aware of the needs of older persons and persons with disabilities” [9].

### **Social Aspects of Aging Population**

The increasing aging of the population has revealed a specific and difficult relationship with digital technology, often leading to various fears such as the elderly-specific *technophobia* which represents the main obstacle to the efficient use of ICT. The aging process involves factors such as health, socio-economic, psychosocial that might have an important influence on the senior and his family.

**Health factors:** The natural aging process comes with continuous degenerative disorders like hearing, vision, perceptual, cognitive and sensory impairments, which often make difficult the use of digital technology. For instance, texts may become difficult to be read, pictures difficult to be interpreted, cognitive problems can cause difficulties of understanding, or mobility constraints can limit the accessibility to ICT.

**Socio-economic factors:** Socio-economics factors are very important in terms of the quality of elderly’s life. The aging generations have slower adapted to the enhanced presence of digital technology, with increased vulnerabilities as consequence.

Technologies that encourage a successful aging in place have the potential to bring major benefits and reduce the social and health costs. It is therefore necessary to reconsider older people as a class of relevant users and provide them appropriate technology as a real support for physical independence, a dignified and integrated life in society.

**Psychosocial factors:** Throughout the process of aging, the man’s personality is changing, reflecting the ability of the psyche to withstand the life’s challenges. Some events may precipitate psychological aging: retirement, children’s departure, the partner's death, and adaptation to the new life might have major psychological implications. Seniors that continue to live at home have minor depression compared to the institutionalized ones.

### **Particularities of Designing User Interfaces for Aging people**

Designing web applications adapted to aging population implies taking into consideration some key principles. The most relevant ones are presented as follows:

#### ***Principle 1: Personalized perceptible information***

An increased degree of acceptability of web applications cannot be achieved without an aging user-centered process.

The interface must effectively communicate the information to the users, regardless his/her physical and cognitive skills. Representation of essential information must be provided under different modes (visual, verbal, tactile). Compatibility between the elderly with a diversity of impairments and digital technology people requires a special attention.

#### ***Principle 2: Multidisciplinary teams for development***

The development process of an aging-friendly web application should be performed by effective multidisciplinary teams [9]. The analysis, design and implementing activity must imply ICT specialists, psychologists, sociologists, usability experts, and elderly users.

#### ***Principle 3: User-centered equability use***

The focus on the user, the targets of the activity, the work domain, the context of use, the user's goals, tasks and needs should guide the development of a web application [7]. Designers should understand the context in which it will be used, including the particularities of the end users and the way in which they interact with the information. Thus, the usage aspects will be centered on the aging user from the early stages of development.

#### ***Principle 4: Reliability, accessibility and flexibility***

The development of the aging-friendly web application should be both iterative and incremental [7]. It is necessary to consider the user's specific needs and allow him to choose among different methods of accessing the interface and the context of use, depending on his/her knowledge, perception and adaptability degree.

#### ***Principle 5: Elderly participation in the design process***

Representative elderly users should actively, early and continuously participate throughout the entire development process and throughout the system lifecycle [7, 9, 13].

#### ***Principle 6: Simple, intuitive design***

The design of the user interface must be done so it could be easily understood by end users and all other stakeholders [11]. The using of the interface must be easy to understand, regardless of the elderly user experience, knowledge, and cognitive skills. Ensuring compliance with user expectations and abilities must be achieved taking into consideration the particular elderly capabilities.

#### ***Principle 7: Anticipating user needs***

Aging people being the target users, the design team should take into account all the aspects that may occur in the context of future use, including the continuous diminishing of vision, hearing, cognitive, and motor abilities.

#### **A Perspective-based Technique for Usability Inspection**

Usability inspection has proved to be a major support for improving the quality of the software applications, even during the development stage. In usability inspection methods, specialists examine usability-related aspects of a system interface by using their experience and knowledge

of usability standards [1]. The most used methods are the following:

- *Heuristic evaluation* - specialists in usability evaluate the interface as concerns the way it follows usability heuristics [8].
- *Cognitive walkthrough* - task-oriented method involving the analysis of an interface by an analyst who goes through its functionalities, simulating step-by-step user behaviour for agreed-upon tasks [8].
- *Feature inspection* - inspectors are given use cases and end results and they test each feature for its availability, understandability and other usability features [6].
- *Pluralistic walkthrough* - a version of cognitive walkthrough which involves not only experts, but also users and software developers, in going through the system and discussing about the dialogue elements [15].
- *Perspective-based inspection*: divides the large variety of usability issues along different perspectives and focuses each inspection session on one perspective [18].

The perspective-based inspection aims:

- “To devise a usability inspection technique by inspecting from one particular perspective at a time.
- To understand the feasibility, effectiveness, and scope of perspective-based usability inspection” [5]

### CASE STUDIES

The objective of the present case studies was to evaluate the usability of two websites addressing aging in place by taking into consideration the degenerative problems of an elderly user. Another target was to identify the perspectives of improving the websites with the help of Big Data.

#### Aging in Place Website

*Aging in Place* is an American website that provides resources and support for an appropriate daily living and home-remodelling in order to ensure a safe and independent life for a getting older person. (see Figure 1) [http://aginginplace.com/]

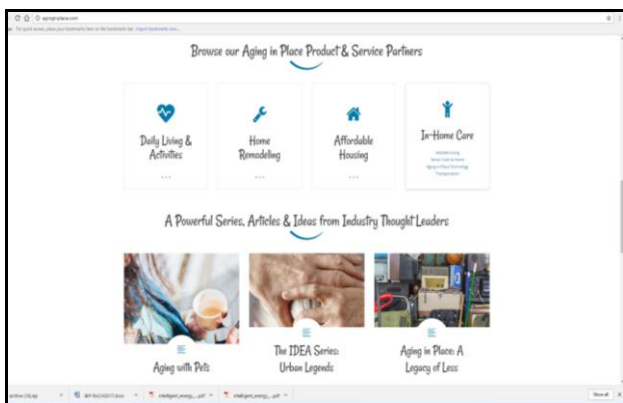


Figure 1: Aging in Place website

#### Ana Aslan International Foundation Website

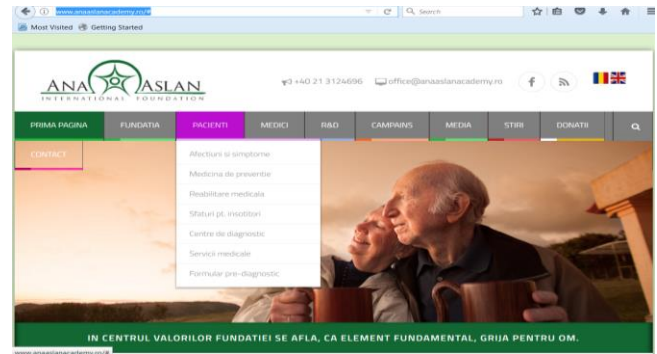


Figure 2: Ana Aslan International Foundation website

*Ana Aslan International Foundation* is a Romanian organization aiming “to integrate scientific progress into the original concept of 3P Medicine - Preventive, Predictive and Personalized medicine -, thus offering patients, medical and scientific bodies, instruments with which to make out of the Medicine of Aging”. The foundation website presents products and services designed to promote longevity, prevent pathological cerebral aging and to facilitate a viable aging in place. (see Figure 2) [http://www.anaaslanacademy.ro/#]

#### Procedure

Perspective-based usability inspection of the websites took place independently, the inspectors being asked to perform specific tasks and qualitative analyses by adopting different perspectives.

1. Aging related perspectives have been defined, taking into consideration the common objective of the two websites, namely providing support for a successful aging in place.
2. The inspectors have received updated information about elderly user experience issues and usability inspection.
3. Specific evaluation tasks and criteria have been defined.
4. The results of the independent inspections have been discussed and prioritized.

#### Usability inspection completion

At the beginning of the usability inspection process, the specific tasks and two papers with examples of perspective-based usability inspection (as references) have been provided to the selected inspectors (see Table 1).

No.	Task
1	Getting information about aging-in-place
2	Getting information about technology for aging-in-place
3	Obtaining information from specialists

Table 1: The evaluation tasks

In this study, 5 inspectors examined independently the two websites and reported the usability problems found in each one for each task:

1. *Expert in application inspector* – testing issues of efficiency or aspects of the website to support the user.
2. *Inspector with disabilities* – three elderly persons with specific disabilities and noting where the product is not accessible.
3. *Psychologist inspector* – looking for violations of psychological principles related to memory, learning, attention, fatigue, etc.

In Table 2 are presented 10 usability factors considered in the QUIM (Quality in Use Integrated Map) consolidated model for usability measurement [19], as well as an 11th factor – *acceptability* - that introduces new criteria specifically for the elderly population: safety, discretion, dependability, non-obtrusiveness, appropriateness, understandability, and trustworthiness [8]. Taking into consideration these criteria, for each usability problem detected in the websites, each inspector recorded information concerning: context, cause, suggestions for improvement, and severity (*severe* – failure to accomplish the task goal, *moderate* – with important impact on task execution but passable and *minor* – with not important impact on task goal).

**Usability inspection results**

The collaborative inspection of usability conducted to a total of 40 usability problems for both websites. Each of the problems detected by the inspectors has been analyzed in order to agree on the severity. The usability evaluation results are presented in Table 3.

Relationship between Factors and Criteria (sub-categories)											
Criteria	Factors										
	Efficiency	Effectiveness	Satisfaction	Productivity	Learnability	Safety	Trustfulness	Accessibility	Universality	Usefulness	Acceptability
Likeability			+								+
Controllability							+	+	+	+	+
Simplicity					+			+	+		+
Privacy							+		+	+	+
Security						+	+		+		+
Familiarity					+		+				+
...											
Safety											+
Discretion											+
Dependability											+
Non-obtrusiveness											+
Appropriateness											+
Understandability											+
Trustworthiness											+

**Table 2: The set of perspective usability inspection criteria**

Task	Aging in place website				Ana Aslan Foundation website			
	Total	Severe	Moderate	Minor	Total	Severe	Moderate	Minor
1	14	5	8	1	11	4	5	2
2	4	1	3	-	4	1	3	-
3	3	2	1	-	4	3	1	-
Total	21	8	12	1	19	8	9	2

**Table 3: Usability problems per task and severity**

Two of the severe problems are related to the lack of specific visual and audio accessibility characteristics for elderly people (e.g. fix dimension of characters in the texts) in both websites and the third is related to Errors management issues in *Ana Aslan Foundation* website (pages under construction).

Most of the moderate problems are related to:

- Difficulties in accepting the information related to aging in place for *Ana Aslan Foundation* website (unstructured and insufficient information).
- Insufficient available information in *Ana Aslan Foundation* website that affects the website usefulness.
- Absence of searching facilities for finding information in *Aging in place* website, even if, in general, elderly users preferred to navigate the website by browsing rather than searching.
- Limitation of obtaining information from specialists for *Ana Aslan Foundation* website, a discussion forum not being included.
- Provision of indications not personalized for the elderly / caregiver and lack of downloadable forms for recording information in both websites.
- Poor clarity of menus, insufficiently structured web pages, lack of consistency of presentation for different pages, all these leading to the reduction of the efficiency of use in both websites.
- The absence of a "Help" button in both websites.

The perspective-based inspection of usability proved to be useful for training the inspectors that are not experts in the application domain, the detection rate of usability problems being higher for them.

**ISSUES FOR FURTHER DEVELOPMENT OF AGING IN PLACE WEB APPLICATIONS**

**Opportunities brought by Big Data**

The last advances in ageing-in-place are focused on incorporating software programs and monitoring device into senior homes, which will help them to increase well-being, provide additional safety nets, and identify any behavioral abnormalities that might indicate a problem [3]. As part of aging in place, a more efficient long-term care is based on the progresses in ICT aimed at developing home monitoring devices and software that allows seniors to live more independently and provides the opportunity of online medical diagnostics, and interventions. Such examples are wearable devices for continuously monitoring senior physical activities and tracking vital signs allowing patients to receive real time help in case of any abnormalities detected in their information, low-cost laptop computers with sensors to read vital signs, and sending data and images to physicians at other locations, integrated platforms for providing inclusive health services to older people, use

of mobile phones and Internet as elements for rehabilitation, education, mentoring, goal-setting, personal feedback, and counseling [2].

All these new devices that can connect elderly people to friends, family, and caregivers without requiring constant, in-person supervision generate a huge amount of information, structured or unstructured. That is why, Big Data is used to facilitate aging in place and one should promote systematic data gathering and analysis in order to provide support to seniors across the continuum of care. Aging in place is a comprehensive domain including public health, medical treatment, nursing care, and welfare work (Figure 3).

The use of Big Data in a collaborative system also brings challenges in terms of information sharing, standardized data gathering, and security of personal information [16].

User Interfaces (UI) help presenting complex Big Data in a visually appealing manner to users. Thus, it has become an integral part of Big Data. “This ubiquitous use of UI among all sections of users has added to its worth as an important player in the big data revolution” [10].

*Ana Aslan Foundation* website creates a bridge to the inclusion of Big Data due to the embedded medical forms suitable for further development and broader access to health systems based on Big Data Analytics.

### General Recommendations

Although the two presented websites are focused on elderly’s health related problems, distinctive functionalities, provided resources and approach make the difference in terms of implementing a practical support for the empowerment of the seniors that chose to age in place.

Thus, *Aging in place* website mainly presents informative resources, while *Ana Aslan Foundation* website also brings forward a pre-diagnostic evaluation form which, together with the several under-deployment functions, represent a step towards developing an interactive ehealth solution.

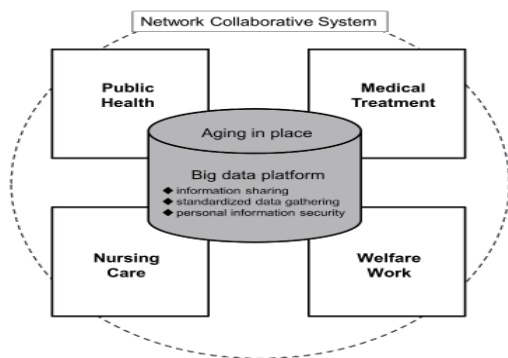


Figure 3: Use of Big Data in a network collaborative system for ageing in place [16]

In the same time, there are some common usability challenges that have to be taken into consideration to better meet the specific necessities and demands of an older user.

Also considering the guidelines specified by Nielsen Norman Group in 2013, in the second edition of “Senior Citizens (Ages 65 and Older) on the Web” [14], some recommendations can be drawn for making the two websites more accessible and engaging for elderly:

- Navigation controls should be avoided for the elderly. Especially seniors with cognitive and motor dysfunctionalities might have difficulties to elaborate mental models of the structure of the websites.
- Homepages, a permanent and visible home button on the screen in both layout and content areas should be very accessible and distinctive.
- The use of the keyboard has to be diminished and, if possible, replaced by a virtual keyboard adapted to the type of content.
- The text and style should be adapted to the visual impairment.
- The terminology used in the user interface should be appropriate and familiar to the one of the elderly.
- The search function should have a design and usage that facilitate a quick and simple finding of information.
- The content should help the seniors to process information and to enhance the understanding of it.

Because aging in place implies the use of digital technology (like Internet of Things) that generate massive amount of health data, Big Data Analytics can properly address the challenge to aggregate and integrate it. This is the reason for the necessity of further development of web applications like the two discussed in this paper with integrated models based on Big Data. Moreover, the health data regarding a senior which are collected from residential homes should be in real-time correlated with the personal ones from the National Integrated Health Information System for a quick and appropriate action in case of an emergency.

For a faster adaptation to the increased ageing of the population, a roadmap with critical information for the development and implementation of best practices, solutions and future directions of eHealth domain in Romania, enlarged with guidelines stipulating design principles for providing appropriate digital services for elderly represents a practical support for an active aging in place.

### CONCLUSION

The global aging at a growing rate, the enhanced everyday presence of ICT and the increased seniors’ technology acceptance and expertise have imposed to reconsider seniors as a class of relevant users that are looking for digital technology able to provide a real support for a dignified and active life in a personal environment. That is



why there is a need to use age-related aspects in designing interfaces based on familiar patterns specific to seniors' degree of understanding and acceptance.

Usability testing should allow the ICT specialists to better meet the requirements of the elderly.

Big Data support aging in place by allowing access to huge amounts of health data and analytical results and by creating a framework for implementing pro-active health and social models.

The results of the presented case studies and the general recommendations can be used both for designing new aging-related websites and for a future development of the two ongoing websites.

Addressing the age-related challenges can bring extended possibilities to better engage elderly with web applications that can support them to improve the management of their health condition and the quality of an independent and active life, namely for a successful aging in place.

#### ACKNOWLEDGMENTS

This work has been developed in the framework of "Using Big Data Technologies in Government Computer Systems" and "ICT in healthcare: A comparative analysis of the eHealth solutions in Member States and successful models in the world" projects funded by the Romanian Ministry of Communication and Information Society (2015-2017).

#### REFERENCES

1. Alexandru, C.A. Usability Testing and Improvement of Telemedicine Websites, Master's thesis, *University of Edinburgh, School of Informatics*, 2010
2. Beard, J.R. and Bloom, D.E. Towards a comprehensive public health response to population ageing. *Lancet*, 385, (2015), 658 – 661
3. Blitchok, A. *Technology For Aging in Place*. <https://changingaging.org/supporting-each-other/technology-for-aging-in-place/>
4. Centers For Disease Control And Prevention, *Healthy Places Terminology*, 2013 <https://www.cdc.gov/healthyplaces/terminology.htm>
5. Experimental Software Engineering Group, *Usability Evaluation*, 2005 <http://www.usabilityhome.com/FramedLi.htm?http://www.cs.umd.edu/projects/SoftEng/ESEG/topics/usability.html>
6. *Feature Inspection* [Internet], Available from: <http://www.usabilityhome.com/FramedLi.htm?FeatureI.htm>
7. Gould, J. D., Boies, S. J. and Ukelson, J. How to Design Usable Systems. In M. Helander, T. K. Landauer and P. Prabhu (eds) *Handbook of Human-Computer Interaction*, Amsterdam: *Elsevier Science B.V.*, 1997
8. Holzinger A. Usability engineering for software developers. *Communications of the ACM*, 48 (1), (2005) 71–74
9. ISO/IEC Guide 71: 2014, <https://www.iso.org/obp/ui/#iso:std:iso-iec:guide:71:ed-2:v1:en>
10. Jon, C. *Importance of UI and UX in Big Data*. (2016). <https://dzone.com/articles/importance-of-ui-and-ux-in-big-data>.
11. Kyng, M., Making Representations Work. *Communication of the ACM*, 38 (1995), 46 – 55
12. Nathan Brown: *Healthcare Data Growth: An Exponential Problem*, 2015 <http://www.nextech.com/blog/healthcare-data-growth-an-exponential-problem>
13. Nielsen, J., Iterative user-interface design, *IEEE Computer*, 26 (1993), 32 – 41
14. Pernice, K., Estes, J., Nielsen, J. *Senior Citizens (Ages 65 and older) on the Web*. Fremont, CA: *Nielsen Norman Group*, 2013
15. *Pluralistic Walkthrough* [Internet], Available from: <http://www.usabilityhome.com/FramedLi.htm?PlurWalk.htm>
16. Song, P. and Yu, C. Public policy response, aging in place, and big data platforms: Creating an effective collaborative system to cope with aging of the population. *Bioscience trends* 9, no. 1 (2015), 1 – 6
17. *Usability Evaluation Methods* [Internet], Available from: <http://www.usabilityhome.com/>
18. Zhang, Zhijun, Victor Basili, and Ben Shneiderman. "Perspective-based usability inspection: An empirical validation of efficacy." *Empirical Software Engineering* 4.1 (1999), pp 50
19. Seffah, A., Donyaee, M., Kline, R.B. and Padda, H.K., Usability measurement and metrics: A consolidated model. *Software Quality Journal*, 14 (2), (2006), 159 – 178
20. Holzinger, A., Searle, G., Kleinberger, T., Seffah, A. and Javahery, H., July. Investigating usability metrics for the design and development of applications for the elderly. In *International Conference on Computers for Handicapped Persons*, Springer Berlin Heidelberg, (2008), 98 – 105

# A user-centric approach to test automation of web-based applications

**Ionuț PETRE**

ICI Bucharest

8-10 Averescu Avenue

ionut.petre@ici.ro

**Dragoș-Marian SMADA**

ICI Bucharest

8-10 Averescu Avenue

dragos.smada@ici.ro

**Radu-Marius BONCEA**

ICI Bucharest

8-10 Averescu Avenue

radu@rotld.ro

## ABSTRACT

Modern society is imposing a faster rhythm in most of the life aspects, with software development being in the front line. Business requirements are stressing development teams for faster releases of new features and functionalities of the applications. The software developers adapted to the market demands by migrating from traditional development models to what is known as “continuous delivery”. This implies short cycles of software development to ensure a higher degree of reliability for application releases. The use of this approach leads to a frequent building, testing and releasing of the software. Modern web applications offer a wide palette of user interactions and reached the level where their behavior is very similar to native applications. As a result, automation of the tests on the app behavior on the client side should be a priority for software development teams. For web-based apps, the client-side experience is the most important factor, as it is the most important factor deciding the application success. Frequent releases can induce many errors in the app’s source code and can affect the functioning on the user side; tests automation is the solution that can prevent flaws in the app behavior at the user.

## Author Keywords

continuous delivery; user-centric test; automatic web tests; app behavior tests; continuous deployment

## ACM Classification Keywords

Software and its engineering

## INTRODUCTION

In a fast-paced society in which technologies are emerging continuously, software development is a continuous task for every business or institution that wants to maintain a high level of quality. Organizations need unprecedented velocity and agility to seize new opportunities and software development teams are under pressure to deliver high-quality software at a faster rate than ever before. Drawn-out development projects are no longer acceptable in the conditions of the modern business demands. The development team must stand up to this challenge while facing the most diverse technology ecosystem in the history of computing [5]. The creation of software is a team effort, with a wide range of aspects that require integration and conflict solving. The number of web-based applications is

increasing, and so are the levels of reliability and complexity required to develop the apps. Scalable, instantly-accessible multi-user, multi-component, multi-tier systems are being required to perform mission-critical tasks in an environment of continuous delivery and operation, with an expectation of zero downtime and easy maintenance and administration [3]. A poor communication between team members or a poor planning of tasks can lead to major failures. In global software engineering, communication and coordination become more challenging, and that fact affects the quality of the product [1].

A proper test automation solution reduces the time and resources required to adequately test software, while also increasing the level of quality of the application under test [2].

## CONTINUOUS SOFTWARE DELIVERY

Continuous delivery represents a software engineering approach in which teams produce software in short cycles, ensuring that the software can be reliably released at any time [4]. The goal is to build, test, and release software application versions at a frequent rate. Continuous delivery is different than the concept of continuous deployment, where any change is automatically deployed to the production environment. Continuous delivery means that the team ensures every change can be deployed to production but may choose not to do it, usually due to business reasons. In order to do continuous deployment, one must be doing continuous delivery [10]. Continuous delivery and tests automation are two symbiotic concepts with the goal of providing a faster release of new versions of the application.

## Agile development

Facing business demands to accelerate time for software release, developers are switching from traditional development models to continuous delivery based on Agile development methods. **Agile development** is appealed to be an innovative and receptive effort to address users’ needs concentrated on the prerequisite to distribute applicable working business applications faster and inexpensive [6]. Through this approach, software is developed in short cycles, thereby ensuring reliability for timely releases. This results in building, testing and releasing the software faster and more frequently. The approach has proven to reduce cost, time and the risk of delivering critical changes to production, thereby allowing

incremental updates to the production system [7]. For development teams to take application agility and efficiency to the level required by the modern society, three major concepts must be followed [5]:

- Focus on modern software architecture paradigms and frameworks. These are the new trend in software development and can reduce considerably the time to launch an application.
- If possible, replace heavyweight platforms with lightweight application infrastructure. This can bring scalability to the system while maintaining fast development.
- Invest in automation and strive for increased agility across the life cycle. Automation is a necessity due to the fact that modern architecture and lightweight infrastructure are characterized by changes which cannot be manually managed.

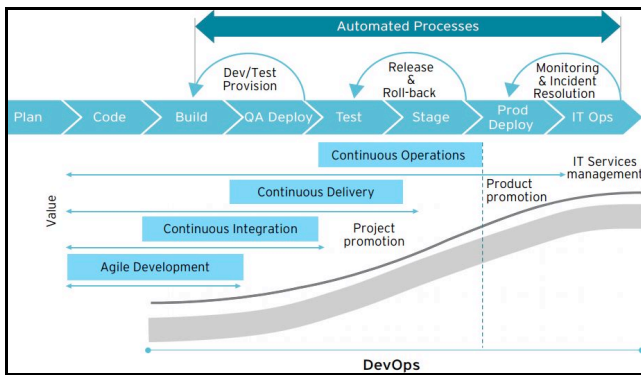


Figure 1 - Continuous Delivery[7]

Most organizations follow a software development cycle plan to develop, maintain, modify or enhance their software. The development cycle is a complex job that allows team managers to estimate development costs, to set timeframes and plan tasks. Developing quality applications require teamwork and complex planning of development cycle. Given the volume of work and tasks that must be performed by the development team, the appearance of flaws can hardly be avoided.

Continuous delivery is driven by continuous integration, where the code is logged-in several times a day and then re-compiled, leading to frequent deployments, which mandates greater levels of process automation. For a successful project, automation of the build and deployment process is critical; this ensures a self-testing build and makes the entire process transparent and agile system [7].

#### Basic considerations on user-centric tests

The web-based applications are written to run in a browser, therefore testing these applications in an automated manner must involve the simulation of browser behavior. The testing automation from a user-centric perspective involves the usage of software tools that perform tests repeatedly on

the application by simulating the actions of an operator or user in a browser. The automation tests for a web-based application must be focused on the user experience, having the goal to raise the users' confidence in the quality of the application. The user-centric approach also provides advantages for the development team and for the clients. The developers can identify and solve issues, and the clients can verify that the application fulfills specifications and requirements.

Test automation is not appropriate in all cases of software development. There are times when manual testing is more suitable to test an application – such is the case when the application interface is changing frequently. Therefore the user-centric automation is the solution when the interface is stable and provides key elements that are rarely or never changed. The alteration of the interface implies the reconstruction of test-cases and an analysis on costs-benefits must be done by the organization prior to the decision of automation.

#### USE CASE OF USER-CENTRIC TEST AUTOMATION Solutions and environments

The use case presented in the article represents the test automation of a complex web application used by the operators at the Romanian Top Level Domain registry. The authentication in the application is done by username and password, with users being already set.

The web application functions over a middleware architecture, meaning that this provides ways to connect the various software components into an application so that they can exchange information with relatively easy-to-use mechanisms. Middleware deals with component communication modes and can be used in a wide range of domains. The application is written in *PHP* while the middleware is written in *Python*. The middleware provides a set of commands that can be integrated into the test scripts. The application is under continuous development and integration with a stable interface, therefore, tests automation was considered a proper solution for the development team.

**Selenium** is an open-source solution consisting of a suite of tools for automating the behavior of web browsers across various platforms. It provides support for several programming languages, such as Java, Python, C#, Ruby. It has two options for usage [8]:

- Selenium IDE – an add-on for Firefox browser to record-and-playback interactions with the browser;
- Selenium WebDriver – a collection of language-specific bindings, to create browser-based regression automation suites and tests. The scripts can be used in different environments.

**Selenium Python** bindings provide a simple API to write functional/acceptance tests using Selenium WebDriver.

Through Selenium Python API the functionalities of Selenium WebDriver can be accessed for widely-used browsers such as Firefox, Internet Explorer, Chrome, Safari, phantomjs, etc [8].

### Tests automation implementation

The starting point is represented by the elaboration of a document containing all the operations that the user can perform on the app. The test assertions in the document were derived from the OASIS Test Assertions Guidelines Version 1.0 [9] that describes best practices in applying a general test assertion model, what test assertions are, their benefits, and how they are created. The main purpose of the test assertions is to remove the ambiguities or statements that can lead to impractical or unnecessary development efforts. The document can be used whenever there is the need for a test change or for creating new tests. If not developed by the specification authors, test assertions should be reviewed and approved by them which will improve both the quality and time-to-deployment of the specification. Therefore, best results are achieved when assertions are developed in parallel with the specification. An alternative is to have the leader of the team that is creating test suites to write the test assertions as well and to provide feedback to the specification authors [9]. In this case, the document was elaborated by the authors of the application specifications in collaboration with the team of operators. The document is updated through the test development process with situations and test cases that the operators missed

**Selenium IDE** comes with an interface for developing automated tests, with a recording feature which records user actions as they are performed. Each user action is translated into a Selenium command that can be personalized. The recording can be exported as a reusable script written in a programming language that can be later executed on a server. Selenium IDE was the tool used for translating into user behavior on the browser for the assertion tests specified in the document. Each test recorded user actions in various key elements – authentication, application menu, links, buttons, or confirmation/warning messages. The records were exported as Python2 WebDriver script files. The recording was done in Firefox browser, on a Windows platform.

The script files of tests must run headlessly on a continuous integration server, running with Linux Centos 6, which does not provide a display output to launch a browser. The issue is that Selenium requires simulating a web browser in order to run the python files exported from IDE. The solution to this problem was **Xvfb**, an X server for machines without hardware display which emulates a virtual framebuffer using the virtual memory of the machine and in our case is used to simulate the display needed for the browser. To link the tests from scripts to Firefox, another tool is required: **Geckodriver**. This, in fact, is a proxy for the W3C WebDriver-compatible clients to interact with Firefox.

The commands for the installation of Selenium, Xvfb, and Geckodriver on the application' server:

```
> yum install Xvfb firefox
> pip2.7 install -U selenium==2.53.2
> wget geckodriver-v0.11.0-linux64.tar.gz
> tar -xvzf geckodriver-v0.11.0-linux64.tar.gz
```

The work procedure followed by the team was to perform the test in Selenium IDE, installed in Firefox. The assertions were made for elements considered stable on the web page, such as <h> elements, action buttons, menu links, alert messages, confirmation messages, user-interaction alert windows:

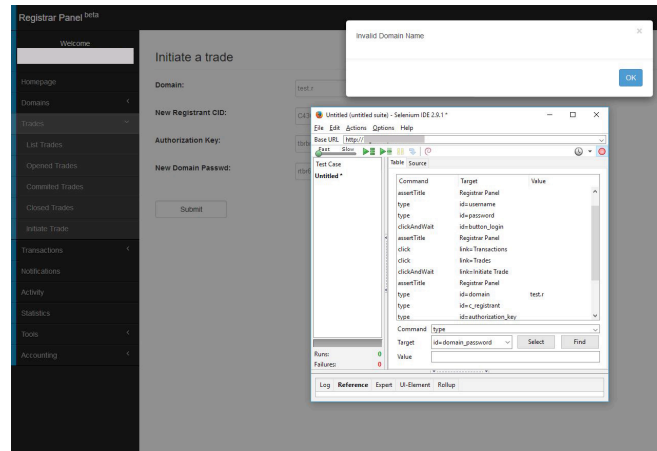


Figure 2 - Use of Selenium IDE for recording web activity

After recording the operator actions in the web application, we run the test from Selenium IDE to ensure proper assertion of the saved record.

Each test is exported as Python2 unit test WebDriver in a CentOS platform. Running the test under Python led to the appearance of errors due to the insufficient transition time between the page and any messages or alerts that appeared. This required the manual setting of additional pauses of 1 or 2 second between the launching and closing of any messages that require user interaction.

### Tests automation server

For a complete automation process, the tests must run without user assistance. **Jenkins**, a Java-based open source solution to deliver continuous build, is the tool to complete this task. It has the capability to monitor any job defined as a cron, SVN or GIT. A continuous integration server is designed to automatically or manually trigger complex workflows to build, test, and deploy software components. Although it is a platform focused on building software systems, Jenkins-CI can easily be expanded with over 800 extensions for complex computational tasks. To create, organize and discuss, the development team uses Atlassian Stash, a Git repository management solution for enterprise

teams. It allows everyone to easily collaborate on your Git repositories. To ensure code stability, good collaboration between developers and fast release cycles, the developers integrated Jenkins to build tests automatically on every pull requests made on the Stash Server.

Stash Webhook was configured on the Stash server to trigger Jenkins automatically. In the workflow, a developer creates a branch on his local repository, completes its task, then commits and pushes the changes to Stash and creates a pull request. When the pull request is merged, Jenkins is triggered automatically to make the build and run the Selenium tests.

### Disadvantages

The development and maintenance of tests automation imply considerable effort on the development team and costs on the client when it comes to complex web applications, especially when the user interface changes frequently. In these cases, it can prove a hard task to create and maintain automated tests for dynamic contents.

The assertion document must be elaborated considering all the aspects, including different account types in case the displayed content or the client interface is different. These situations require additional effort.

### CONCLUSION

The decision on whether to perform automated tests varies from one organization to another, but in times where Agile development is spreading for faster software development, the automation of tests becomes a requirement for a successful implementation.

User-centric tests automation presents major advantages:

- Better experience for the users on the client side, as the automation increases the application's reliability;
- Reduces the load on the development team. Code review is performed faster as the reviewers only verify the code that passes the client-side tests;
- Faster integration of new developers in a team. The tests assertions document leads to a better understanding of the application for new-comers;
- In the event of committing bad code, the committed code fails the tests, is automatically denied so the code flaws do not reach the production environment

Besides the development of the tests, there are virtual machines that need to be provisioned and configured, and testing environments must be set and deployed. But in a long-term vision, tests automation reduces the development

time and decrease the risk of bugs that might interfere with the application behavior and raises the users' confidence.

### ACKNOWLEDGMENTS

This work has been supported by a Romanian grant, financed by ANCS under COGNOTIC 1609 0802 / 2016. We thank all the colleagues who participated in the tests automation development and also provided helpful comments on previous versions of this document.

### REFERENCES

1. DeFranco, J.F., Laplante, P.A. An Integrative Review and Analysis of Software Development Team Communication Research. *IEEE Transactions on Professional Communication PP(99)* · June 2017
2. Dustin, E., Rashka, J., Paul, J. Automated Software Testing: Introduction, Management, and Performance, 1999
3. Hearn, J.E., The Development of an Automated Testing Framework for Data-Driven Testing Utilizing the UML Testing Profile, 2016
4. Lianping, C. "Continuous Delivery: Overcoming Adoption Obstacles", Continuous Software Evolution and Delivery (CSED) *IEEE/ACM International Workshop on*, pp. 84-84, 2016
5. Knoernschild, K., Knipp, E., Watson, R., Kenefick, S., Brian, D., Olliffe, G., Holz, B., Dayley, B., Cheetham, L., Modernize Application Development to Succeed as a Digital Business, *Gartner*, March 30, 2016  
<https://www.gartner.com/doc/3270018?refval=&pcp=mp>
6. Saleh, S.M., Rahman, A.M., Asgor, K.A. Comparative Study on the Software Methodologies for Effective Software Development. *International Journal of Scientific & Engineering Research*, Volume 8, Issue 4, April-2017
7. Sitaraman, S., Bar, R., Test Automation Strategies in a Continuous Delivery Ecosystem, Cognizant, 2016
8. <http://www.seleniumhq.org/>, accessed in May 2017
9. Test Assertions Guidelines Version 1.0, 2013  
<http://docs.oasis-open.org/tag/guidelines/v1.0/cn02/guidelines-v1.0-cn02.pdf>
10. <https://martinfowler.com/bliki/ContinuousDelivery.html>, accessed in May 2017



# Predicting User Activities using GPS Collections

**Simona Andries**

Faculty of Computer Science,  
“Alexandru Ion Cuza” University of Iasi  
General Berthelot, No. 16  
simona.andries@info.uaic.ro

**Adrian Iftene**

Faculty of Computer Science,  
“Alexandru Ion Cuza” University of Iasi  
General Berthelot, No. 16  
adiftene@info.uaic.ro

## ABSTRACT

The purpose of this paper is to retrieve as much information as possible from a set of GPS coordinates. The information obtained and presented in this study are the most frequented places of a user and the modes of transport used to get from one place to another. The considered modes of transport are: walking, biking and travelling by car, train or airplane. This kind of information can be used in many types of applications like recommendation systems or it can be studied in order to have a better understanding of the human activity and find new ways to improve it.

## Author Keywords

Prediction; DBScan; GPS.

## ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI):  
Miscellaneous.

## General Terms

Human Factors; Design; Measurement.

## INTRODUCTION

The everyday schedule of people gets more and more busy, but their desire to be up to date with the activities of their friends and family doesn't change. Another great factor that should be considered is that since the free time is so short, it is important not to waste it in traffic, searching for an address in a new neighborhood or having unpleasant experiences in a restaurant. Consequently, the need to always improve applications that register and manage routes or recommendation systems applications appears. This type of application can be used by people with an active life or by the companies that need to manage the routes made by the employees in a day.

The GPS coordinates of a user represent an extremely important source of information, which utilized in a good way can bring a lot of benefits. For example, with the help of a recommendation system like Google Traffic [6], the user can choose a route to a certain location that is longer than the usual route, but that doesn't have traffic jams and this way a lot of time is saved. Another example is that if during the city holidays the traffic is busier than usual in certain areas of the city, it can be inferred that the next year, in the same period, the traffic will be busy also. Based on this information a recommendation can be made that could help a lot of users.

Nowadays, mobile phone and especially smartphones are something extremely common among people. These phones are equipped with GPS receptors, they provide wireless access to internet or 3G/4G services and this makes social network applications available to users at any moment of the day.

Intelligent Transport Systems [8] represent a suite of technologies meant to bring improvements to the different modes of transport and how they are managed. They range from navigation systems, traffic lights control systems, automatic plate number recognition and highway surveillance cameras to parking sensors.

One of the main technologies that is part of Intelligent Transport Systems is Global Positioning System (GPS). According to [3], the GPS functionality is based on using space satellites as reference points for localizing points on earth. By using a very precise measurement of distance in a straight line between the receptor and at least four satellites, the position (latitude, longitude, altitude) of any point of earth can be determined. The distance between the satellite and the receptor [3] is calculated by timing the period needed by the radio signal to get from the satellite to the receptor. The data registered by GPS equipped devices will be studied in order to determine the user's locations of interest and to find the modes of transportations used to get from one place to another.

The application proposed in this paper represents a way to facilitate the way GPS trajectories logged by devices can be analyzed. Based on these trajectories information of interest can be determined that are important for the user, for the companies that monitor the activities performed by employees on the field and last but not the least to public institutions interested in improving traffic in cities. Having this kind of information available, they can take better decisions or consolidate the opinions made by using other kinds of observations or studies.

The main information that can be obtained based on GPS trajectories are: the locations of interest of a user and the amount of time spent there, the modes of transport used to get from one place to another and the routes that are most commonly used by a user. In this paper the first two points will be discussed the most.

## SIMILAR APPLICATIONS

In the last years there were developed more and more applications that help the users to be connected with their close ones or recommendation system applications that offer suggestions based on preferences like Geolife.

Geolife [5] a Microsoft project, is a social network based on locations that helps users connect and share between them their experiences. This application analyses the location history of the users in order to determine the most popular routes and locations among them with the purpose of creating recommendations for trips or everyday routes. Based on this information and also on the circles a person is present in, the application measures the similarity between users and makes personalized recommendations (see Figure 1).

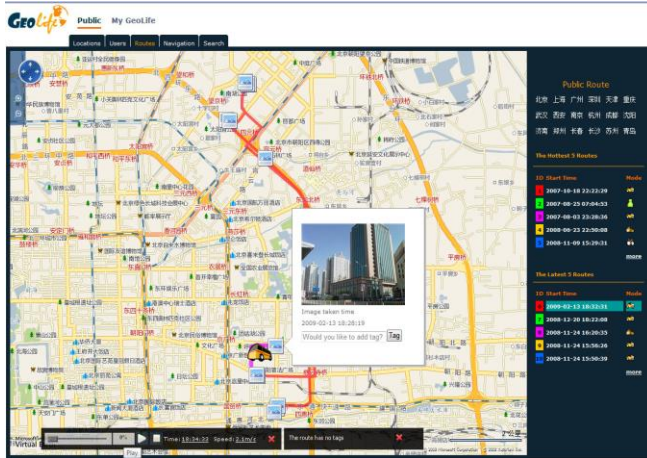


Figure 1. GeoLife application [10]

By analyzing the location history using Data Mining techniques, Geolife determines the most important locations for a user or his frequent travelling routes. Other functionalities offered by this application are the suggestion of locations of interest based on the recommendations made by other users or to suggest itineraries that can be used while travelling. A common problem encountered while developing recommendation systems is to find a reliable way to calculate the similarity between users using their location history and deciding the level of interest a certain location will have for a user.

According to [19], another common difficulty encountered in this research area is to find the mode of transport of a user. People can change frequently the mode of transport in a day and some common problems are given by the always changing traffic conditions from big cities or by the poor quality of the GPS coordinates registered by devices sometimes.

There are a number of studies that have the purpose of determining the mode of transport used on a route based on a set of GPS coordinates. The majority of them identify the segments covered by foot, bike, car and train and are mainly based on the speed calculated between consecutive

points. In addition to these studies [2, 14 and 15], this paper brings as a novelty the possibility to determine segments travelled with the airplane. Another difference is that in order to correctly determine the routes travelled by train the true north direction is used.

The studies that identify also the public modes of transport use GIS data or other kinds of algorithms like the one proposed in the study [18]. This algorithm is based on a group of segments that are labeled with the mode of transport organized in a graph as a base. An edge represents the route of the user from a point to another and for each edge a weight is calculated. This weight represents the distribution of probabilities for the possible modes of transport and each new segment is classified based on the previously computed probability for the edge corresponding to the segment. This method has the disadvantage that it requires a very big initial set of data in order to classify new segments with a good precision.

GPS Trackmaker [7] is an application that allows the upload of files containing GPS logs with the purpose of displaying a map with the travelled routes and computes some metrics based on it like acceleration, average speed and maximum speed.

## GPS DATA ANALYSIS

In this study a set of GPS Tracks from Microsoft Research Asia Geolife project [4] was analyzed. These GPS tracks [19] were collected from 182 users over a period of five years. The majority of the coordinates are registered in Beijing, but also in other cities from China. This set of data consists of a sequence of GPS coordinates accompanied by time stamps. As specified in the user manual [19] of the GPS tracks, the trajectories were logged using different types of devices equipped with GPS receptors and have different recording times ranging in the interval 1-5 seconds or 5-10 meters. The data set encompasses a variety of movements of the users, ranging from the routine home – work and back to sports activities or relaxation.

The data set is analyzed with the purpose of extracting new information from a big quantity of data by summarizing the information in order to make it easier to analyze. Based on the information gathered from this analysis statistics can be created to determine the most popular mode of transport, the busiest time of the day in a certain place and many others.

Simultaneously, this data set can be used in a wide area of research like location recommendation or to study user trajectories. In this paper it was used to determine the locations of interest, the amount of time spent there and the modes of transport used to get from one place to another by the user.

The paper Applying Data Mining in Prediction and Classification of Urban Traffic [11] proposes a way to classify data in order to understand and predict traffic based

on decision trees. A decision tree has the structure of a flowchart. In this kind of structure each internal node is tested for a certain property and each branch of a node represents the result of a test. The leaves of the tree represent a class and the main node is called a root. The resulting tree can be used for classifying information because each new record passes through the root node and after a test it is passed on to the next node. This process goes on until a leaf node is reached that represents the class (decision) and the path to the leaf node is a set of rules.

In this paper decision trees were used to classify the segments resulted from the splitting of the trajectory based on certain criteria. The attributes of the segments that were used in the classification are: speed, maximum speed, altitude, total time and the distance of the segment. As previously stated, the tree leaf node represents the decision and it can have one of the values: walk, cycle, car, train, and airplane.

### Mining Important User Places

One of the most important things that can be discovered based on the GPS coordinates are the most important locations for the user, where he spends the most time, or stops by regularly for a shorter or longer period of time. Figure 2 presents a basic routine covered by a user in a day.

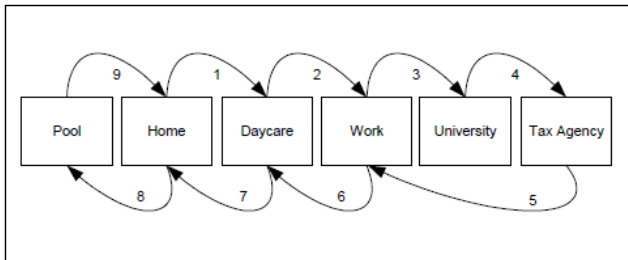


Figure 2. Typical places visited by a person in a day [Zhou, Bhatnagar]

*How can we determine these stops?* An idea would be to consider a stop a place where the speed is 0 km/h for a minimum period of time. This approach has limitations because GPS devices can have signal shortages, or the speed could be close to 0km/h because the user is stuck in a traffic jam. Moreover, a stop is not so easy to determine, because even if the user is standing still, the captured GPS logs are not 100% precise and can have noises [22].

Other approaches use more sophisticated techniques like splitting a journey in trajectories. This way, a 5 minute break indicates a stop, whereas a 5 minute trajectory at a very low speed indicates that the car is stuck in a traffic jam. Andrienko [1] and Zheng [21] consider in their studies that a location is important if a person spends a lot of time there. Stops are considered as consecutive GPS coordinates such that the distance between the points doesn't exceed a certain threshold (e.g. 50 meters) and the time spent there exceeds a minimum threshold (e.g. 30 minutes).

CB-SMoT [12] considers that a stop is determined based on the variation of speed on a trajectory. This way, the stops are the segments where the speed is lower than the average speed of the trajectory. A stop during a trip indicates that we got to a location of interest or close to one. An interesting area of trajectory study is determining and labeling the points of interest (POI) that motivate the stop and, if none is found, labeling the stop with "unknown".

### Determining POI for an User

In order to determine the POI for a user 3 clustering algorithms were studied in order to find the most suitable one. The studied algorithms are: K-means, DBSCAN and DJ-Clustering [16 and 22]. The density based clustering algorithms like DJ-Clustering and DBSCAN need two variables: Eps (distance between nodes) and MinPts (minimum number of points needed to form a cluster). These values must be chosen according to the density of the registered GPS logs.

The main idea of the DJ-Clustering algorithm is to group close objects and represent them using the minimum bounding rectangle. The data structure used to represent the information is an R-Tree where each node stores two things: a way to identify the children nodes and the bounding box of the children nodes. In order to have a better understanding of how this algorithm works a library was used to notice the results. The Figure 3 shows the result of applying the algorithm on the data set.

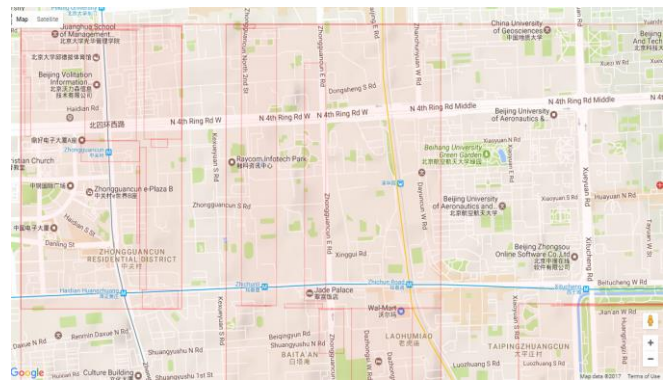


Figure 3. The result of the application of the DJ-Clustering algorithm

After analyzing the three algorithms, the conclusion was that the algorithm suitable for resolving the problem of determining the points of interest is DBSCAN. DJ-Clustering is more useful for resolving other kind of problems like finding all the locations of interest around a certain point and K-Means clustering has the disadvantage that it requires as input the number of clusters that need to be obtained, and this number can vary a lot depending on the size of the data set and the frequency of the GPS coordinates or the period of time over which the coordinates were registered. In our case, if the period is very short, it should be possible to get 0 clusters, so setting a predefined number of clusters is not helpful in this case.

In Figure 4 can be noticed a couple of locations of interest determined for a user using DBSCAN clustering algorithm and some of the tracks of a user from a week. Each line color represents a different day.

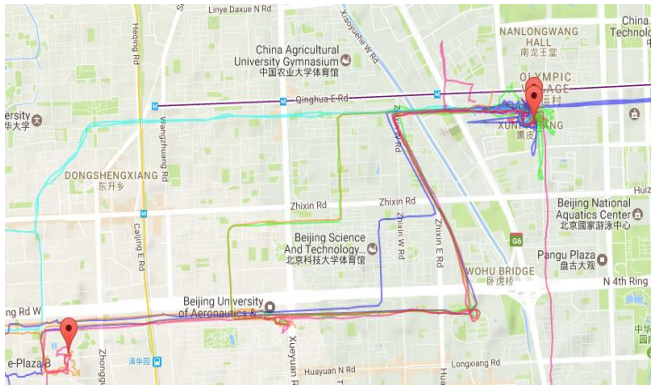


Figure 4. Point of interests for User 052, 14 - 20 July 2008

### Determining Segments for a User

The concept of trajectory is associated to the always changing position of an object over a period of time, so a trajectory is spatio temporal concept. The concept of journey means that the movement of an object has the purpose of travelling from one point to another. A journey take a finite amount of time and covers a distance in space (SPACCAPIETRA et al., 2008).

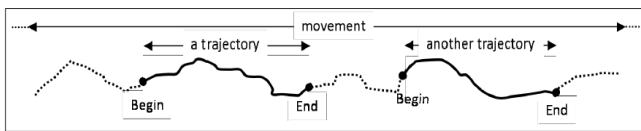


Figure 5. Trajectories extracted from a list of GPS coordinates [13]

A lot of applications are not interested in keeping and analyzing big volumes of data, but only the segments they are interested in. Figure 5 represents with interrupted line a section from the GPS logs of a person and with a full line two sections that were identified as relevant. Each trajectory is defined using spatio temporal coordinates called Begin and End.

Among the information that can be determined based on GPS coordinates there is the possibility to find common features of trajectories. This is based on evaluating similarities that allow defining classes of trajectories. The most important techniques that are based on machine learning are classification and cluster analysis. In this paper, a trajectory is formed from a sequence of GPS coordinates that represent the trip made by a user in order to get from one point to another and can be covered using multiple modes of transport. A trajectory is then split in segments, each segment having associated the transportation mode that was used.

Yu Zheng, in the paper Learning Transportation Mode from Raw GPS Data for Geographic Applications on the Web

[20] considers two premises in order to do the segmentation of the GPS coordinates. The first one is that the change of transportation mode must be done by walking and the second one is that between segments there must be a break or a stop where the speed is very close to 0 km/h.

The segmentation of GPS coordinates takes place in two steps and it also has a preprocessing and post processing step. The preprocessing has the purpose of reducing the errors that could appear because of wrong GPS coordinates. The preprocessing algorithm is based on the distances computed relative to five points (A, B, C, D, and E). For each of them a sum of the distances to the other four points is made. If the sum reported to C is the biggest, the coordinates of point C are replaced by the point resulted by calculating the midpoint between C and the point given by the minimum between the segments (AC, BC, CD, CE).

The first step is to split the journey in trajectories (separated by longer stops) and the second step is to split the trajectories in segments corresponding to the different means of transportation used. The segmentation process leaves from the premise that every person switches multiple modes of transport while travelling and this is why the segments are split in two types: walking and non-walking. Then, the non-walking segments are classified using the decision tree.

The split of the journey in trajectories is made determining the places where the user spends at least 30 minutes. If such a place is found, it means that the previous point is the end of a trajectory and the next point represents the beginning of a new trajectory. If no such place is found it means that the whole set of data represents a single trajectory.

One of the difficulties encountered in the segmentation process is given by what wrongly seems to be a very frequent change or transportation mode that is impossible. This is caused by the poor traffic conditions or by the stops at traffic lights. To solve this issue, at the end of the classification a post processing is made to merge the segments that have a very short duration.

### Determining the means of transport used

One way of segmenting the trajectories is using the means of transport used by a person. This information is useful in planning the public transport system. This type of segmentation is based on the main features of each transport means that can be used. Usually, a walker has a maximum speed of 5 km/h, while using a bicycle a person can reach around 10 to 15 km/h and the buses have clear paths.

Based on speed, acceleration, and gear change rate, and using a technique such as supervised learning or decision trees, we can determine for each trajectory the means of transport used. This information is important because, for many people, the route to the office involves walking, one or more means of public transport and walking again. A



library was used to build the decision tree, and the algorithm that classifies segments based on the built-in tree was written for this application.

Identifying the means of public transport used is not trivial because many aspects need to be considered. Not just the use of recorded speed is sufficient, as traffic conditions can cause unusual variations, but a machine may have a speed similar to that recorded by a bicycle. And often, people often use more means to get from one point to another.

One important thing to keep in mind is that most public transports have unique features that can help us identify the means of transport used in a particular segment of a journey. A common point for all means of transport is the average speed of each. For this work, we have set some intervals specific to each transport mode based on the data set available. Because some users have annotated their trajectories with very important information, the means of transport used over time, we used this information to deduce the typical speeds at the locations where the GPS coordinates were recorded. Thus, we determined all the GPS coordinates specific to the annotated time interval, and we calculated the average speed for each record based on them.

Some means of transport are easier to identify than others because they have some defining features that eliminate any other possible variation. The *plane* is the easiest to identify because of the flying altitude. The cruising altitude for a line jet is about 10,000 m. And another important feature is the speed of about 850 km/h.

The next means of transport that can be relatively easily identified is *walking*. This is because speed is very small and constant throughout the route. The range used in this work for the average speed is 2-7 km/h.

The *bicycle* is the next transport vehicle analyzed. Beginning with this means of transport, precise identification of the means of transport becomes very difficult, and this is due to the fact that there may be periods of time when the user seems to walk or stop when he is actually at the traffic light. The average speed considered for this means of transport is in the range of 7 to 20 km/h.

The means of transport that can be found among users' annotations are *car*, *bus*, *taxi*. These were treated in this paper as one, namely *car*. In order to identify the means of public transport, it is necessary to use GIS data, which is not part of the scope of this paper, but which may be an improvement that is worth considering for the future.

The next mode of transport is the *car*, which typically has average speeds between 20 and 120 km/h. This means of transport is the most complex because it can "borrow" characteristics from the means of transport described above in situations that are very common. These situations are stopping and starting from traffic lights and last but not least traffic jams. Very often, track sequences such as car-

stop-car or car-bicycle-car or combinations of these two have a very short duration for the intermediate route. Starting from a logical premise, namely that a person can not have such a route in real life, we have combined these sequences from the paths into one. This step of combining the segments of the route was done at a later stage of their classification.

The *train* is the last means of transport analyzed in this paper. The minimum speed considered for this means of transport is 120 km/h and the maximum speed is 250 km/h. As anyone can tell, the train and the car have many similarities due to the wide range of journeys that can be made using these means of transport. Therefore, the determination of this means of transport takes place in two steps. The first takes place in the segment classification using a decision tree, this classification is based on the average and maximum speed, the distance traveled and the duration of the route. Practically, the annotated results following classification may restrict the type of trains determined on high-speed or interurban trains. The second step occurs after segment classification and is based on direction. Bearing [17] is the angle between the North Pole and the trajectory line of a moving object. What distinguishes the car from the train is the angle with which it can make curves. A car can at any time make a curve that has an angle of 90 degrees, while a train can make this curve, but much wider and at a greater distance. In this step, the direction is analyzed, and if there is no sudden change of direction, the labeling of a segment previously tagged with a car on the train changes.

## CASE STUDIES

Due to the presence in the data set of tracers accompanied by labels, we have the possibility to evaluate the accuracy of the results obtained in this paper. Here are some examples of the data set.

### Use Case 1 – Train (User 010, 28 March 2008)

This case aims to analyze the accuracy with which trajectories traveled by train are identified. Figure 6 shows the user-labeled segments, and Figure 7 shows the result obtained by the proposed application.

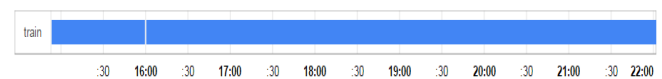


Figure 6. Representing data labeled by the user (Use Case 1)

For this user, there are periods where no GPS coordinates have been recorded, which are labeled automatically with missing data. As can be seen from Figures 6 and 7, a difference from the data labeled by the user is given by the identified stops. It is very likely that these are stations where the train has stopped, and the proximity of a walk towards the stop can be an extra argument for this hypothesis.



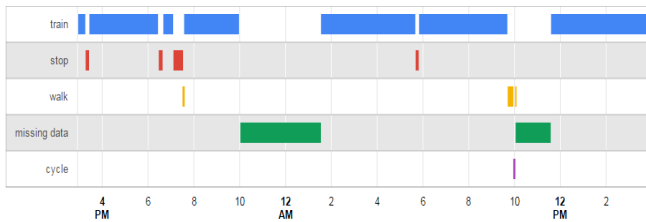


Figure 7. The result obtained by the application (Use Case 1)

**Use Case 2 – Airplane (User 128, 30 September 2008)**

The purpose of this case is to demonstrate the accuracy with which the airplane is identified. As can be seen in Figures 8 and 9, the airplane is identified with very good accuracy. The same can be said about the train. We can also notice that some segments have been mistakenly labeled with the bike because of the speed too low (see Figure 10) with which the vehicle is moving and that there were generally high traffic fluctuations in the city.

As in the previous example, we can see labeled segments with missing data, the lack of GPS coordinates being a common problem.

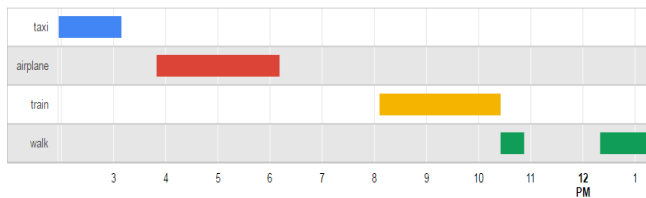


Figure 8. Representing data labeled by the user (Use Case 2)

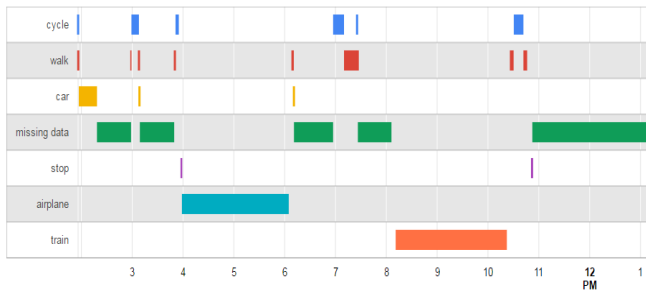


Figure 9. The result obtained by the application (Use Case 2)

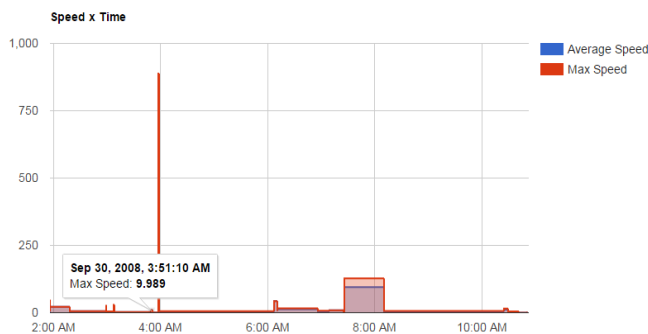


Figure 10. Speed variation over time

**Use Case 3 – Walk-car-walk-car (User 179, 21 August 2009)**

The purpose of this case is to illustrate the behavior of the application for the alternative walk-car-walk-car segments.

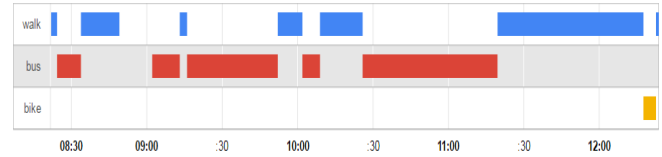


Figure 11. Representing data labeled by the user (Use Case 3)

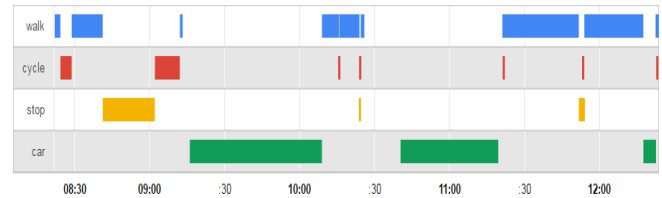


Figure 12. The result obtained by the application (Use Case 3)

As you can see in Figures 11 and 12, the first two walk segments have been identified with good precision. The first bus segment was misidentified as a bicycle because of its very low speed (14 km/h). We can also see that a walk segment is not identified, but is embedded in a car segment. Among the segments labeled by the user, one can also see a labeled bike that was not correctly identified due to the very high speed recorded (38 km/h).

Although the segment identification algorithm has a pre-processing stage in which noise has been reduced, it may not be completely reduced. The third and fourth cycle segments were identified differently due to high speeds. At the same time, it is possible for the user not to explicitly label all means of transport. He could, for example, run when he wants to catch the green color at a traffic light, thus influencing the outcome of the algorithm. Running is not part of the possible means of transport, users who tagged the segments did not have this option.

**Use Case 4 – Bicycle (User 020, 28 October 2011)**

This case aims at analyzing the accuracy with which the bicycle routes are identified and analyzing a variant that would improve the accuracy of the determination of the means of transport used.

Figure 13 shows a segment labeled by the user that has been wholly completed by bicycle. As can be seen, Figure 14 shows two variants of classification of segments traversed. The first variant uses the maximum speed of all the speeds calculated between each two points. The second variant uses the maximum speed identified based on an index ( $0.95 \cdot \text{number of points}$ ) between determined speeds (ordering upwards), thus ignoring the noise points or at least part of them. For the case studies presented above, it is worth mentioning that there is no difference between segment labels in these two cases.

An alternative to the 95th percentile could be the use of an algorithm that eliminates noises such as the Kalman Filter [9] that estimates and corrects the coordinate position.



Figure 13. Representing data labeled by the user (Use Case 4)

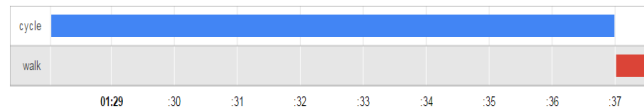
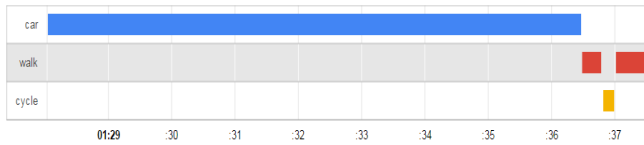


Figure 14. The result obtained by the application before (up) and after (down) application of 95th percentile (Use Case 4)

## CONCLUSION

From a large amount of information, we had originally obtained some easy-to-read information by anyone. For good viewing, interest locations have been marked on a map, and the result of segment classification has been displayed in a graph. As we saw, we used a clustering algorithm to determine where the user spends the most time. One of the limitations of clustering algorithms is that there is no exact method to determine if a cluster is correct. The user of the algorithm must decide whether the data resulting from the clustering satisfies the problem it wants to solve. Abnormal data obtained by applying clustering algorithms may correspond to data of major interest, or may not be representative and removed prior to clustering or post-processing.

One of the advantages of the proposed solution is that it can be easily improved in the future by adding new classification criteria to the decision tree. One of the criteria that may be added is acceleration, and adding new classification criteria improves the precision with which segments are classified. Also, in order to get a better result, it would be worth analyzing to see if there are algorithms that completely remove the noise from the dataset.

We also look into the future and based on the locations of interest determined and the routes used to get from one point to another one can build a chart that could be used to analyze frequently visited routes by the user.

## ACKNOWLEDGMENT

This work is partially supported by POC-A1-A1.2.3-G-2015 program, as part of the PrivateSky project (P\_40\_371/13/01.09.2016).

## REFERENCES

1. Andrienko, N., Andrienko, G. Designing visual analytics methods for massive collections of movement. *Cartographica*, 42(2), (2007), pp. 117-138.
2. Biljecki, F., Ledoux, H., Van Oosterom, P. Transportation mode-based segmentation and classification of movement trajectories. *International Journal of Geographical Information Science*, 27(2), (2013), 385-407.
3. El-Rabbany, A. *Introduction to GPS: The Global Positioning System*. Artech House, Boston, London. www.artechhouse.com, 194 pages, (2002).
4. GeoLife GPS Trajectories: <http://research.microsoft.com/en-us/downloads/b16d359d-d164-469e-9fd4-daa38f2b2e13/> (Accessed last time on 1 June 2017)
5. Geolife: <http://research.microsoft.com/en-us/projects/geolife/> (Accessed last time on 1 June 2017)
6. Google Traffic: <https://www.google.ro/maps> (Accessed last time on 1 June 2017)
7. GPS Trackmaker: <http://www.trackmaker.com> (Accessed last time on 1 June 2017)
8. Intelligent Transport Systems: <http://www.etsi.org/technologies-clusters/technologies/intelligent-transport> (Accessed last time on 1 June 2017)
9. Kalman, R.E. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*. 82 (1), (1960), 35-45.
10. Microsoft UseCase: <https://www.microsoft.com/en-us/research/project/geolife-building-social-networks-using-human-location-history/> (Accessed last time on 1 June 2017)
11. Nejad, S. K., Seifi, F., Ahmadi, H., Seifi, N. Applying Data Mining in Prediction and Classification of Urban Traffic. *World Congress on Computer Science and Information Engineering*, (2009).
12. Palma, A. T., Bogorny, V., Kuijpers, B., Alvares, L. O. A clustering-based approach for discovering interesting places in trajectories. In Proceedings of the 2008 ACM symposium on Applied computing (SAC '08). ACM, New York, NY, USA, (2008), 863-868.
13. Parent, C., Spaccapietra, S., Renso, C., Andrienko, G., Andrienko, N., Bogorny, V., Damiani, M. L., Gkoulalas-Divanis, A., Macedo, J., Pelekis, N., Theodoridis, Y., Yan, Z. Semantic trajectories modeling and analysis. *ACM Comput. Surv.* 45, 4, Article 42 (2013), 32 pages.
14. Schuessler, N., Axhausen, K. W. *Processing GPS Raw Data without Additional Information*. 15 pages, (2008).
15. Stopher, P., Clifford, E., Zhang, J., Fitzgerald, C. Deducing Mode and Purpose from GPS Data. *Working paper ITLS-WP-08-06*. Institute of Transport and

- Logistic Studies, the Australian Key Center in Transport and Logistic Management, the University of Sydney, (2008).
16. Tan, P. N., Steinbach, M., Kumar, V. *Introduction to Data Mining*. Addison – Wesley, (2006).
  17. Rutstrum, C. The Wilderness Route Finder. *University of Minnesota Press*, (2000).
  18. Zheng, Y., Chen, Y., Li, Q., Xie, X., Ma, W. Y. Understanding transportation modes based on GPS data for web applications. *ACM Trans. Web* 4, 1, Article 1 (January 2010), 36 pages, (2010).
  19. Zheng, Y., Fu, F. Xie, X., Ma, W. Y., Li, Q. *Geolife GPS trajectory dataset - User Guide*, 26 July 2011.
  20. Zheng, Y., Liu, L., Wang, L., Xie, X. Learning transportation mode from raw gps data for geographic applications on the web. In *Proceedings of the 17th international conference on World Wide Web (WWW '08)*. ACM, New York, NY, USA, (2008), 247-256.
  21. Zheng, Y., Zhou, X. *Computing with Spatial Trajectories*. Springer, 2011.
  22. Zhou, C., Bhatnagar, N., Shekhar, S., Terveen, L. Mining Personally Important Places from GPS Tracks. In *Data Engineering Workshop, 2007 IEEE 23rd International Conference on*, Istanbul, Turkey, (2007).

# A case study of usability evaluation: the Center for Active Ageing website

Marilena Ianculescu<sup>1</sup>, Ovidiu Bica<sup>1</sup>, Dragos Daniel Iordache<sup>1</sup>, Costin Pribeanu<sup>1,2</sup>

<sup>1</sup>National Institute for Research and Development in Informatics - ICI Bucharest

Bd. Maresal Averescu, nr.8-10, Bucharest, Romania  
{manina, ovi, iordache, pribeanu}@ici.ro

<sup>2</sup>Academy of Romanian Scientists

Splaiul Independentei nr.54, Bucharest, Romania

## ABSTRACT

The population is ageing and the percentage of the elderly is continuously growing. A key priority to address the demographic challenges is to implement more efficient e-services. The websites for active aging provide useful information and advice to old people, informal caregivers, and medical staff. In order to be an effective help, these websites should be evaluated for usability. The objective of this paper is to present a case study in evaluating a website aiming to promote the active aging. The evaluation results revealed several important usability problems that have to be fixed in the next version. Most of the important usability problems are related to the information architecture, navigation, compatibility with the user and task, task guidance, and support.

## Keywords

Usability, usability inspection, usability heuristics, an aging population, active ageing.

## INTRODUCTION

This millennium is characterized by a fast growth of the elderly in the overall population (see Figure 1). Causes explaining the phenomenon of ageing include lowering birth, medical progress, and rising living standards; these factors have increased the number of the elderly by reducing morbidity and mortality.

The population is aging and the percentage of the elderly is continuously growing. Active aging is a concept stating that the participation in various activities in the later life contributes to health and wellbeing. Several factors exist that influence a healthy and active aging [1].

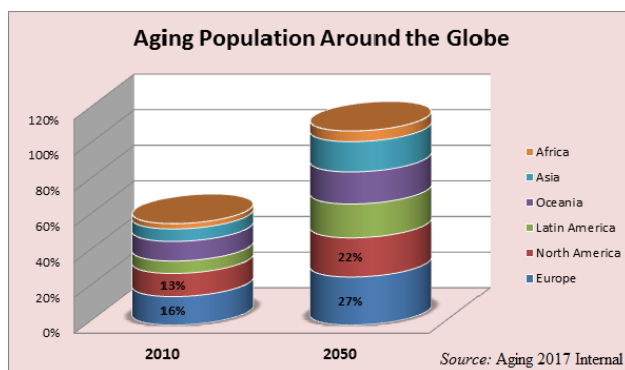


Figure 1. Projected aged population in 2050

A key priority to meet the demographic ageing challenges is to implement more efficient e-services that are able to support integrated solutions for all the actors involved in the elderly's healthcare.

This needs more investment and innovation to adapt and rethink the care systems, in particular through a better integration of services and continuity of care, so as to avoid as much as possible the institutionalization. As pointed out in the European vision, "digital innovation ... has the potential to ensure that the ageing population remain independent and active in society, can receive coordinated care and enjoy living longer in their homes" [17]

The websites for active ageing are providing useful information and advice to old people, informal caregivers, and medical staff.

Designing for web usability in the area of active aging is not easy. Despite the numerous projects at European and national levels, the research on usability in this area is too narrowly focused. It seems that in most cases there are too little resources for usability studies, therefore, few case studies exist, and, consequently, little usability data is available. Another issue is related to the diverse levels of education and web literacy of the target users. This leads to consider specific design requirements [1, 3, 9,10].

This paper presents a case study of usability evaluation of the Center for Active Ageing website that was developed during a national research project. The website provides useful information for the elderly as regards the medical services, home adaptation recommendations, and monitoring of the main physiological parameters.

## CASE STUDY

### The Center for Active Ageing website

PROlonging ACTIVE life for an independent and healthy Ageing (ProActive Ageing) is a project funded by the PNCDI II national research program. The partnership is composed of Clinical Hospital "Dr. Victor Babes", University of Medicine and Pharmacy "Carol Davila", "Ana Aslan" National Institute of Gerontology and Geriatrics, National Institute for R&D in Informatics and Siveco S.A.

The main outcome of the project is the ProActive Ageing integrated platform for online service, structured into three thematic modules:

- Center for Active Ageing.

- Comprehensive learning courses for specialists taking care of ageing people.
- A Toolkit to Support Medical Research Addressed to Healthy and Active Ageing.

More details about the project could be found in [6]

The Center for Active Ageing aims to support an independent life of elderly, providing a comprehensive approach to a successful ageing in place, health management, and an increased health literacy.

The purpose of the center is to assist and encourage people to better understand the ageing process and to facilitate the access to information about the main aspects of growing older. It is committed to providing opportunities to the elderly, their formal and informal caretakers to take advantage of using an eHealth solution in order to better monitor and manage the everyday life and activities of an ageing person.



Figure 2 Center for Active Ageing

The Center for Active Ageing provides information and awareness of the degenerative age-related processes, with an accent on lifelong learning focused on ageing persons that want to remain active, healthy and independent.

The ultimate goal of the Center for Active Ageing is to promote the independence of aging people by providing a holistic approach in order to maintain an independent life as long as possible.

The website is structured into the following components:

- Home care
  - Adjustments to the housing needs of the elderly;
  - How to employ a caretaker for an elderly person;
  - Types of home care services.
- Active and independent person
  - Improvement of personal knowledge about healthy ageing;
  - Aspects of an active and independent ageing;
  - Means of maintaining an active life;
  - Self-assessment tests
- Useful information
  - Emergency (Ambulances, Hospitals Foundations, Palliative Care, Physicians, NGOs)
  - Legislation concerning elderly.

The target beneficiaries of the Pro-Active Ageing Centre are the elderly that wish to preserve an active and healthy life, preferably in their own house, their caretakers, and families. Because ageing is a process that starts around 30, another type of user is any person that wants to prevent and delay the age-related health disorders.

The information and data stored can be displayed without any restriction by any user of the module.

Only authorized health specialists have the right to introduce/modify/delete information into this module's components. A user is authorized to enter data and information in this module only by following an authorization procedure supervised by the online platform administrator.

### Evaluation approach

Usability evaluation aims at finding, documenting, and reporting usability problems. The goal of the evaluation is to identify usability problems, help the developers to fix the problems and, this way, improve the usability of the interactive system. A usability problem has been defined as any aspect of the user interface which might create difficulties for the user [12]. Integrating the usability evaluation in the software design cycle creates a user-centered-design attitude, ensures a good practice, and provides useful design knowledge and case studies.

Usability evaluation methods could be classified into two main categories: inspection methods (faster and cheaper) and user testing (more expensive) [4]. Usability inspection (also termed as expert evaluation) is done by experts that are testing the user interface with the goal to anticipate usability problems.

There are several inspection methods. The most widely used are the heuristic evaluation [13]. Several authors pointed out that the heuristic evaluation is oriented more towards fault finding than to goal achievement [7, 8]. Rolf Molich et al. [11] demonstrated that the usability evaluation results depend on the selected tasks, methodology and evaluators.

In this study, a set of 25 usability heuristics has been used for documenting the usability problems. The evaluation is task-oriented since the heuristics are used to explain and document each usability problem as well as to train evaluators. In this respect, the method is more similar to the heuristic walkthrough [16] than to the heuristic evaluation.

### Method and procedure

Before starting the evaluation, each evaluator received the evaluation tasks, the set of usability heuristics, and two papers with examples of usability inspection.

Table 1. The evaluation tasks

No.	Task
1	Create a user account on the platform
2	Obtaining information about the monitoring of physiological parameters
3	Getting information about the home adaptation
4	Sending opinions on the medical services at home

In this study, 3 evaluators tested the website independently and reported the usability problems found for each task.



Two evaluators are usability experts and the third is an expert in the application domain. The evaluation tasks are presented in Table 1.

Then the usability problems have been consolidated per task based on the “similar changes” principle [5].

The following information has been recorded for each usability problem: context, anticipated difficulties, cause, suggestions for improvement, usability principle (heuristic) violated, and severity.

The problems have been documented and explained with a set of usability heuristics that are based on the integration of usability principles proposed by Nielsen & Molich [13] with the ergonomic criteria of Bastien & Scapin [2].

The initial set of heuristics has been used in the usability evaluations of municipal websites [14]. Later on, the set of usability heuristics has been refined in order to better address some user-centred design issues[15]. The usability heuristics are presented in Table 2.

Table 2. The set of usability heuristics

User Guidance	
1	Visibility of system status
2	Prompting
3	Immediate feedback
4	Information architecture
5	Grouping / distinction
6	Legibility
Workload	
7	Concision
8	Recognition instead of recall
9	Minimal actions
10	Information density
Adaptability and control	
11	Flexibility and efficiency of use
12	Experience of the user
13	Explicit user actions
14	User control
Error management	
15	Error prevention
16	Quality of error messages
17	Error correction
Consistency and standards	
18	Consistency
19	Compliance with standards and rules
20	Significance of codes
Compatibility	
21	Compatibility with the user
22	Task compatibility
23	Task guidance and support
24	Help and documentation
25	Esthetic design

According to the potential effect on user’s task, the severity has been assessed as severe, moderate and minor. Severe means a failure to accomplish the task goal or a significant loss of data or time. The problem is moderate if it has an important impact on task execution but the user is able to find a way. A minor problem is irritating the user but the impact on the task’s goal is not important.

A good practice requires fixing all important usability problems (severe and moderate) before the first release of the website.

## Usability Evaluation Results

The number of problems detected by each evaluator varied from 15 to 28. Then the individual problems have been analyzed in order to agree on the severity. A total of 12 problems have been discarded (false problems).

The collaborative consolidation resulted in a total of 32 usability problems, of which 27 are important.

The usability evaluation results are presented in Table 3.

Table 3. Usability problems per task and severity

Task	Total	Major	Moderate	Minor
1	8	-	8	-
2	7	-	7	-
3	7	1	6	-
4	10	1	6	3
<b>Total</b>	<b>32</b>	<b>2</b>	<b>27</b>	<b>3</b>

The two major problems are related to the failure to open a document (requirements for a safe home) and to find the web page for sending an opinion.

Most of the moderate problems are related to:

- It is difficult to understand at a glance what is the usefulness of the center of active aging (mixed information, unclear message to the users)
- Profile page displays technical data (WordPress features) that are irrelevant to the user.
- Redundant functions and buttons, redundant web pages (having just one link).
- Indications for monitoring physiological parameters are not personalized to the elderly / caregiver and there are no downloadable forms to record the data.
- Page titles are not clear, web pages are poorly structured, with long sentences.
- The absence of a "Help" button.

The heuristics are useful to analyze the overall usability of an application according to ergonomic criteria. Most of the important usability problems are related to:

- Information architecture: unnecessary web pages, redundant information, some missing web pages;
- Compatibility with the user: long phrases, technical information, information irrelevant for the user;
- Task guidance and support: inadequate explanation, lack of forms to record the physiological parameters;
- The help and documentation: lack of a help menu/button.

The individual detection rate varied between 37.55% and 46.88%. The average detection rate was 30.95%. The agreement between any two evaluators varied between 4.17% and 21.74% (severity has not been considered). The average any two agreement was 10.15%.

The usability heuristics proved to be also useful for training the expert in the application domain. Surprisingly, the detection rate was higher and the lowest agreement was between the two usability experts.

## DISCUSSION AND CONCLUSION

Usability evaluation before the website release is not a common practice. As pointed out in a previous work [15], the web is mainly designed for functionality and much less for usability. It seems that few developers are taking a

user-centered approach that requires knowing the goals a user wants to accomplish.

This case study revealed several usability problems that are important for the users. Although the major problems could be easily fixed, there are several problems that require a redesign of the website in order to improve the information architecture and navigation.

The analysis of results revealed an issue: the website has been designed having in mind three categories of user: elderly, caregivers, and medical staff. This approach is questionable and creates problems both for the developers and the evaluators since the requirements are different. For example, old people and informal caregivers need a simple navigation and content easy to read while the medical staff needs a lot of medical information which is hard to read and understand by other people.

The results are useful for the developers who have to find a solution to the aforementioned issue and to fix the usability problems. The recommendation was to split the application into two parts, one for the elderly and informal caregivers and the second for the medical staff and professional caregivers.

A general recommendation for the developers of websites is to carry on a task analysis before the design. Knowing the tasks a user wants to perform makes it possible to design the information architecture and, most important, to provide an adequate user guidance.

This case study was also useful for evaluators since it revealed some shortcomings as regards the set of usability heuristics. During the collaborative consolidation, it was sometimes difficult to agree on the explanation of usability problems. The conclusion was that it would be useful to provide a definition and several guidelines in order to detail each heuristic and help the evaluators to document each problem. In a future work the heuristics will be refined and both guidelines and examples will be given.

#### Acknowledgement

This work was supported by the Romanian PNCDI II project under 335/2014.

#### REFERENCES

1. Annear, M., Keeling, S., Wilkinson, T., Cushman, G., Gidlow, B., & Hopkins, H. (2014). Environmental influences on healthy and active ageing: A systematic review. *Ageing and Society*, 34(4), 590-622.
2. Bastien, A. & Scapin, D.L. (1993). Ergonomic criteria for the evaluation of human-computer interfaces. *Technical report No. 156, INRIA, Roquencourt, France.*
3. Castilla, D., Garcia-Palacios, A., Miralles, I., Breton-Lopez, J., Parra, E., Rodriguez-Berges, S., & Botella,

- C. (2016). Effect of web navigation style in elderly users. *Computers in Human Behavior*, 55, 909-920.
4. Cockton, G. Lavery, D. Woolrych, A. (2003). Inspection-based evaluation, Jacko, J.A., Sears, A. (Eds.), *The Human-Computer Interaction Handbook*. LEA, 273-292
5. Hornbaek, K. & Frokjaer, E. (2008). Comparison of techniques for matching of usability problem descriptions. *Interacting with Computers* 20, 505-514.
6. Ianculescu, M., & Alexandru, A. (2016). Silver Digital Patient, a New Emerging Stakeholder in Current Healthcare. ProActiveAgeing: A Case Study. *Studies in Informatics and Control*, 25(4), 461-468.
7. Jong, M., Lentz, L. (2006). Scenario evaluation of municipal Web sites: Development and use of an expert-focused evaluation tool. *Government Information Quarterly* 23, 191-206, 2006.
8. Kamper, R. (2002). Extending the usability of heuristics for design and evaluation: Lead, follow and get out of the way. *International Journal of Human-Computer Interaction* 14(3&4), 447-462.
9. Kurniawan, S. H., & Zaphiris, P. (2003). Web health information architecture for older users. *It & Society*, 1(3), 42-63.
10. Miesenberger, K., & Petz, A. (2014). Easy to read on the web—state of the art and research directions. *Procedia Computer Science*, 27, 318-326.
11. Molich, R. Ede, M. Kaasgaard, K. Karyukin, B. (2004) Comparative usability evaluation, *Behaviour & Information Technology* 23(1), 65-74.
12. Nielsen, J. (1993). *Usability Engineering*. Academic Press, New York.
13. Nielsen, J., and Molich, R. (1990). Heuristic evaluation of user interfaces, *Proceedings of ACM CHI'90*, 249-256.
14. Pribeanu C., Marinescu R.D., Iordache D.D., Gheorghe-Moisii, M. (2010). Exploring the usability of municipal websites: A comparison based on expert evaluation results from four case studies, *Informatica Economică*, 14(4), 87-96.
15. Pribeanu, C. (2014). Extending and refining usability heuristics to better address user-centered design issues in the development and evaluation of municipal websites. *Informatica Economica* 18 (1), 83-92.
16. Sears, A. (1997). Heuristic walkthroughs: Finding the problems without the noise. *International Journal of Human-Computer Interaction*, 9(3), 213-234.
17. \*\*\* (2017). Strategic Vision Developed by Stakeholders. *Blueprint Digital Transformation of Health and Care for the Ageing Society*, 6. Available at:  
[https://ec.europa.eu/newsroom/document.cfm?doc\\_id=40787](https://ec.europa.eu/newsroom/document.cfm?doc_id=40787)

# Using expert evaluation to assess the usability of the National Library of Software Programs

Lidia Bajenaru<sup>1</sup>, Alexandru Ionut Marinescu<sup>1</sup>, Costin Pribeanu<sup>1,2</sup>, Dragos Daniel Iordache<sup>1</sup>

<sup>1</sup> National Institute for Research and Development in Informatics - ICI Bucharest

Bd. Maresal Averescu, nr.8-10, Bucharest, Romania  
{bajenaru, ionut, iordache, pribeanu}@ici.ro

<sup>2</sup> Academy of Romanian Scientists

Splaiul Independentei nr.54, Bucharest, Romania

## ABSTRACT

In the last years, there is an increasing demand to develop a national repository for software programs, especially in the area of public administration. This kind of e-service is required both by vendors interested in promoting their products and funding institutions, interested to estimate the costs and to improve the cost/benefit ratio. The National Library of Software Programs (BNP) is a software application that has been developed during a national research project. This paper presents a case study of usability inspection of this web application. The results revealed several usability problems that were mainly related to the difficulty to create an account and to register online a new product, failure to download the information about a product and the lack of information about products.

## Keywords

Usability, usability inspection, usability heuristics, digital libraries, software catalogs.

## INTRODUCTION

Over the past decade, the global software and services market has known a continuous expansion having as growth engine the quantitative and qualitative evolution of user requirements, coupled with the evolution of ICT technologies, each provider trying to impose on the market with innovative, increasingly complex and diversified software products.

In this context of the competition of the available software solutions, it has appeared a need for computer systems to provide facilities for grouping in one place the software vendors along with their products, and to provide for users the right tools for search, retrieve and recommendation of software products complying with their requirements.

At European level, OSOR - Open Source Observatory and Repository for European Public Administrations has been launched with the purpose to support and encourage the collaborative development and reuse of publicly funded applications for European public administrations [14].

In the last years, there is an increasing demand to develop a national repository for software programs, especially in the area of public administration. This national repository is required both by vendors interested in promoting their products and funding institutions, interested to estimate the costs and to improve the cost/benefit ratio.

This paper presents a case study of usability inspection of a web application: National Library of Software Programs (BNP – Biblioteca Nationala de Programe).

BNP is a software catalog that enables the registration of software products and provides useful information about these to interested users. The web application has been developed during a research project and has been evaluated for usability in March 2017.

The rest of this paper is organized as follows. The following section presents some related work in the area of usability evaluation and usability of digital libraries. In section 3, the case study is presented. The paper ends with a conclusion and future work.

## RELATED WORK

### Usability evaluation

The ISO 9241-11 standard defined usability as the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use [7].

The goal of the usability evaluation is to identify usability problems, help the developers to fix the problems and, this way, improve the usability of the interactive system. A usability problem has been defined by Nielsen [12] as any aspect of the user interface which might create difficulties for the user.

The usability problems are rated according to the potential effect on user's task on three severity levels: major, moderate, and minor:

- major: failure to accomplish the task goal or a significant loss of data or time.
- moderate: has an important impact on task execution but the user is able to find a way.
- minor: is irritating the user but the impact on the task's goal is not important

A good practice requires fixing the important usability problems (severe and moderate) before the first release of the application.

There are two main categories of usability evaluation methods: the inspection methods (expert evaluation) and the user testing.

Usability inspection is carried on by experts that are evaluating the user interface against a set of widely accepted usability principles. The inspection methods are less expensive but more subjective (depend on the

evaluator's expertise). Another advantage is that usability inspection can be carried on in the early stages of the development process [4, 20].

The usability inspection provides two kinds of measures:

- Quantitative: number of usability problems in each category.
- Qualitative: detailed description of individual usability problems.

There are a lot of usability inspections, from which the most widespread is the heuristic evaluation proposed by Nielsen & Molich [13]. The evaluation is done against ten usability principles (heuristics).

Another set of principles has been proposed by Bastien & Scapin [3] in the form of ergonomic criteria. The set is structured into six groups: user guidance, workload, adaptability and control, error management, consistency, and compatibility. For each criterion, the definition, the rationale and some examples of guidelines are given.

Since heuristics and ergonomic criteria are too general, a lot of research has been devoted to extending them with more specific or more detailed usability principles.

The literature review of Quinones & Rusu highlighted several ways of developing and validating heuristics in order to address the evaluation requirements of specific application domains [18]. They noticed that the most frequent approaches to developing heuristics are either based on existing usability heuristics or on a methodology to create usability heuristics.

Heuristic evaluation has been criticized because it is mainly oriented towards fault finding than to the task goal [9]. As Cockton et al. [4] pointed out, rule-based methods such as heuristic evaluation or guidelines-based evaluation are system-centric and evaluate the usability as a property of the system.

Based on several comparative evaluations of the same website, Molich et al [11] concluded that the evaluation results depend on the selected tasks, methodology, and evaluators.

An inspection method that is task oriented is the heuristic walkthrough [19]. Another task-based inspection method has been proposed by Pribeanu et al. [16]. In this method, the usability problems are explained with an extended set of usability heuristics. This set integrates the heuristics of Nielsen & Molich with the ergonomic criteria of Bastien & Scapin. The structure follows the six general ergonomic criteria. The method has been used for the evaluation of municipal websites. Later on, the heuristics have been revised in order to better address some specific user-centred design issues [17].

### Usability of digital libraries

The term "digital library" applies mostly to repositories of publications in a digital format that can be downloaded or viewed. By analogy, the catalogs of software products may be included in the same category. Similar usability problems could be found in both digital library systems and software catalogs, regardless the type of content.

A lot of research has been carried on to assess the quality and usability of digital libraries [2, 8]. However, the interest in usability is low and the focus is more on usability as a quality attribute than on the evaluation.

Dalkiran et al. [5] consider usability in the context of digital libraries as "the ease of use, proliferation and the extent of satisfaction it provides to its users". Kling [10] draws attention to usability issues that go beyond the design of interfaces of multi-user systems, such as digital libraries, and to the organizational usability design issues. Pighin & Brajnik [15] evaluated the retrieval techniques which are a typical feature of the software catalogs.

## CASE STUDY

### The BNP web application

The Software National Library (BNP) is a software library developed by ICI Bucharest with the purpose to collect, store, archive, manage and make available the software products developed in Romania [1]. BNP has created as an OSOR-compliant software registration structure so that to enable the information exchange between the national and European platforms.

The entities included in the BNP are defined in a unitary way in a sufficiently flexible format to allow the description of any type of software products. In Figure 1, the home page of the application is presented.

Filea Produs	Acronim	Den Firmă	Denumire Produs	Domeniu	Data Insc
aaaa	AGER Business Tech			Cultura	-
CID	Institutul National de Cercetare-Dezvoltare In Informatica		Centrul Informatonal de Dermatologie	Informatica medicala	14-MAY-15
Zone Tracking	AGER Business Tech		Gestione incidente "DATA ZONE"	Administratia publica centrala sau locala la nivel judet. Fime monitorizare fote auto; Poitie	01-JUN-14
ANPH	Institutul National de Cercetare-Dezvoltare In Informatica		Site destinat persoanelor cu handicap	Incluziune	05-MAR-14
COSTCONS	Institutul National de Cercetare-Dezvoltare In Informatica		Sistem informatic de analiza a costurilor pe ciclu de viata al constructor.	Administratie publica	05-MAR-14
CROS Supply Chain	S.C. CHRISTIAN GAVRILA S.R.L.		CROS Supply Chain	Distributie; Producte	05-MAR-14
Charisma Analyzer	Total/Soft SA		Charisma Analyzer	Administratie publica; Constructii; Distributie; Financiar-Bancar; Medical-farmaceutic; Producte; Proiectare; Retail; Servicii	05-MAR-14

Figure 1. The home page of BNP

The main objectives of BNP are:

- Efficient structuring of information about software products that are used in the public administration institutions and in the business environment;
- Promoting the software products, especially Romanian ones, through a set of modern procedures for information dissemination;
- Increasing the degree of interoperability of information systems from public administration;
- Stimulating the competitiveness of the IT business environment, including exports, also by facilitating the development of partnerships for the development and use of IT products.

The Software National Library (BNP) centralizes in an on-line catalog the existing commercial software products and those developed through research projects. BNP provides

a platform and a flexible development environment that facilitates access to and dissemination of information related to the software products developed in Romania

The potential users of the BNP system are the software developers (private and state-owned companies), software vendors, and the software users (public administration, businesses, academia, and researchers).

A first prototype has been launched at the end of 2016 and several software providers were invited to register their products in the library.

### Method

In this study, an expert evaluation method has been used that is task-oriented. Four experts tested the application in a task-based approach with the purpose of anticipating the difficulties of a real user.

The evaluation tasks are presented in Table 1.

Table 1. The evaluation tasks

No.	Task
1	Create a user account on the platform
2	Obtaining information about the software products
3	Adding a new product on the platform
4	Sending opinions and suggestions

The usability problems are detected and rated following a task-based approach. The usability problems have been explained and documented by using the set of 25 heuristics presented in Table 2 [17].

Table 2. The set of usability heuristics

User Guidance	
1	Visibility of system status
2	Prompting
3	Immediate feedback
4	Information architecture
5	Grouping / distinction
6	Legibility
Workload	
7	Concision
8	Recognition instead of recall
9	Minimal actions
10	Information density
Adaptability and control	
11	Flexibility and efficiency of use
12	Experience of the user
13	Explicit user actions
14	User control
Error management	
15	Error prevention
16	Quality of error messages
17	Error correction
Consistency and standards	
18	Consistency
19	Compliance with standards and rules
20	Significance of codes
Compatibility	
21	Compatibility with the user
22	Task compatibility
23	Task guidance and support
24	Help and documentation
25	Esthetic design

The evaluation has been performed in two steps:

- Individual evaluation: each evaluator tested the application independently and recorded the usability problems for each task.
- Collaborative consolidation: removing the duplicates, removing the false usability problems, agreeing on a list of unique usability problems, agreeing on the severity rate and finalizing the description.

The following information has been recorded for each problem: context and location, anticipated difficulties, cause, suggestions for improvement, usability principle (heuristic) violated, and severity.

The similar usability problems were integrated following the “similar changes” technique [6].

The reliability has been assessed with two indicators: the average detection rate and the average agreement between any two evaluators.

### Evaluation Results and discussion

The number of problems detected by each evaluator varied from 10 to 18. In the second step, the individual problems have been analyzed in order to eliminate the duplicates and the false problems, agree on the severity, and produce a common problem description. A number of 3 problems have been discarded (false problems).

The collaborative consolidation resulted in a total of 28 usability problems, as shown in Table 3. The detection rate varied between 32% and 34% with a mean of 39.29%. The average any two agreement was 20.91%.

Table 3. Usability problems per task and severity

Task	Total	Major	Moderate	Minor
1	9	1	5	3
2	11	2	5	4
3	3	1	2	-
4	5	-	2	3
<b>Total</b>	<b>28</b>	<b>4</b>	<b>14</b>	<b>10</b>

Four major problems have been detected:

- Difficulty to create an account for a new user.
- Failure to download the information about a product.
- Lack of information about all products (the application only enables the display of recently registered products),
- Difficulty in registering online a new product.

The user guidance is very weak. There are important usability problems that are related to the organization of menus. The horizontal menu is placed above the page header which makes it difficult to notice. The vertical menu is weakly structured and makes it difficult to find the information needed.

The main cause is the fact that the application has not been designed for the online registration of software products. The main concern was related to a rigorous verification of the vendor and of the information provided. This added a lot of bureaucracy, including the requirement to send a formal request by mail. This approach proved to be wrong because many software providers lost their interest and gave up to send their products. This is clearly shown by



the relatively small number of products that have been registered.

There is no search facility to find the software products according to a search criterion. In general, it is difficult to find how to register a new product or how to send an opinion. Other moderate usability problems are related to the error prevention: mandatory data fields are not marked and there is no data validation.

The minor usability problems are related to inappropriate annotation of the user registration section, the English version of the application is not functional and the software catalogs present insufficient information.

## CONCLUSION

This case study brought useful results for the developers. Overall, the BNP web application is far from being usable. First priority is fixing the software bugs and prepare the support for the online registration.

It is obvious that the application has not been designed having in mind a set of clearly defined user tasks. The second priority is to define the users' tasks that are critical for the successful use of the application.

The next version should reconsider the design in order to improve the information architecture, provide adequate user guidance, ensure the online registration of products, and include a search engine.

This work was the first step in the usability evaluation of the application. Before the release of the new version, the usability inspection will be repeated. Then a user testing will be conducted in order to finalize the evaluation report and prepare the official launching of the application.

## Acknowledgement

This work has been supported by two Romanian grants, financed by ANCS under COGNOTIC 1609 0101 / 2016 and COGNOTIC 1609 0601 / 2016.

## REFERENCES

1. Bajenaru, L., Marinescu, I. A., Tomescu, M. (2015) Modernizarea sistemului Bibliotecii Naționale de Programe prin utilizarea mediului de dezvoltare Oracle APEX. *Revista Romana de Informatica si Automatica*, 25( 3), 17-28.
2. Balog, A. (2011). Testing a multidimensional and hierarchical quality assessment model for digital libraries. *Studies in Informatics and Control*, 20(3), 233-246.
3. Bastien, A. & Scapin, D.L. (1993) *Ergonomic criteria for the evaluation of human-computer interfaces*. Technical report No. 156, INRIA, Roquencourt, France.
4. Cockton, G. Lavery, D. Woolrych, A. (2003) Inspection-based evaluation, Jacko, J.A., Sears, A. (Eds.), *The Human-Computer Interaction Handbook*. LEA, 273-292
5. Dalkiran, O., Aker, I., Semanur, O., Zehra, T., Sevgi, K.T. (2014) Usability Testing of Digital Libraries: The Experience of EPrints. *Procedia - Social and Behavioral Sciences* 147, 535 – 543.
6. Hornbaek, K. & Frokjaer, E. (2008). Comparison of techniques for matching of usability problem descriptions. *Interacting with Computers* 20, 505-514.
7. ISO 9241-11 (1998) Ergonomic requirements for office work with visual display terminals (VDTs) - Part 11 Guidance on usability.
8. Jeng, J. (2005). What is usability in the context of the digital library and how can it be measured?. *Information technology and libraries*, 24(2), 47-56.
9. Jong, M., Lentz, L. (2006) Scenario evaluation of municipal Web sites: Development and use of an expert-focused evaluation tool. *Government Information Quarterly* 23, 191-206, 2006.
10. Kling, R., Elliott, M. (1994) Digital library design for organizational usability. *SIGOIS Bull*, 15(2), 59-70.
11. Molich, R. Ede, M. Kaasgaard, K. Karyukin, B. (2004) Comparative usability evaluation. *Behaviour & Information Technology* 23(1), 65-74.
12. Nielsen, J. (1993). *Usability Engineering*. Academic Press, New York.
13. Nielsen, J., and Molich, R., 1990. Heuristic evaluation of user interfaces, *Proc. ACM CHI'90*, 249-256.
14. OSOR.eu - Open Source Observatory and Repository for European public administrations. [www.osor.eu](http://www.osor.eu), accessed in 2016.
15. Pighin, M., & Brajnik, G. (2000). A formative evaluation of information retrieval techniques applied to software catalogues. *Journal of Systems and Software*, 52(2), 131-138.
16. Pribeanu C., Marinescu R.D., Iordache D.D., Gheorghe-Moisii, M. (2010) Exploring the usability of municipal websites: A comparison based on expert evaluation results from four case studies, *Informatica Economică Journal*, 14(4), 87-96.
17. Pribeanu C (2014) Extending and refining usability heuristics to better address user-centered design issues in the development and evaluation of municipal websites. *Informatica Economica* 18 (1), 83-92.
18. Quiñones, D., & Rusu, C. (2017). How to Develop Usability Heuristics: A Systematic Literature Review. *Computer Standards & Interfaces* 52, 89-122.
19. Sears, A. (1997). Heuristic walkthroughs: Finding the problems without the noise. *International Journal of Human-Computer Interaction*, 9(3), 213-234.
20. Wharton, C., & Lewis, C. (1994). The role of psychological theory in usability inspection methods. *Usability inspection methods*, 341-350.

# Rhythm analysis in chats using Natural Language Processing

**Irina-Diana Niculescu**  
Politehnica University of Bucharest  
313 Splaiul Independentei,  
Bucharest, Romania  
irinaniculescu93@gmail.com

**Ștefan Trăușan-Matu**  
University Politehnica of Bucharest  
313 Splaiul Independentei,  
Bucharest, Romania  
and  
Research Institute for Artificial Intelligence  
and  
Academy of Romanian Scientists  
stefan.trausan@cs.pub.ro

## ABSTRACT

Rhythm is important both in daily speaking and literary texts, as it is a means of expressing the writer's/speaker's feelings and ideas. A good human-computer interface in natural language, both spoken and written should also pay attention to rhythm. This paper presents an analysis of rhythm in chats, based on the analogy between conversations and literature and the importance of repetition. In the case of chats, rhythm characterizes involvement, which can be an useful indicator of a successful collaborative learning session.

## Author Keywords

rhythm; natural language processing; chats; repetition

## ACM Classification Keywords

I.2.7 Natural Language Processing: Text analysis.

## INTRODUCTION

This paper presents a system for rhythm analysis in chats. It continues the research made in this domain in the past by the same authors, that is the analysis and comparison of rhythm in poetry, prose, and political speech, by means of three approaches of dividing in rhythmic units: Solomon Marcus approach [1], Boychuck et al. approach [2], Mihai Dinu approach [3], the last of them only for poems. Solomon Marcus approach [1] consists of dividing a text in verses/sentences, and Boychuck et al. approach [2] delimits a text according to punctuation marks, coordinate conjunctions and subordinate conjunctions. Mihai Dinu approach [3] divides each verse in metrical units that contain a single stressed word and the unstressed words that precede it.

In the implementation of these approaches, we considered that only the content words from a text are stressed. For each rhythmic unit obtained, the rhythmic factors proposed

by Solomon Marcus [1] are computed: the rhythmic structure, the rhythmic index, the upper and lower rhythmic limit, the rhythmic diameter. Moreover, the rhythmic factors calculated for the entire texts are of utmost importance, as they represent a means of comparing rhythm in different texts.

## STATE OF THE ART

Regarding rhythm in conversations, Deborah Tannen [4] discusses how “repetition, dialogue, and imagery create involvement in discourse, especially conversational discourse” [4, pg 25]. It is analyzed the similarity between the literary texts and conversations, thus “literary writing elaborates strategies that are spontaneous in conversation” [4, pg 30] and “involvement strategies are the basic force in both conversational and literary discourse by means of sound and sense patterns” [4, pg 31]. At the sound level, the participants “become rhythmically involved”, while at the sense level, they are “meaningfully, mythically involved” [4, pg 31].

Deborah Tannen analyzes also the importance of repetition in conversations [4, Ch. 3]. With regard to the link between repetition and rhythm, the repetition of words and phrases from a given text represents the unit of measure of the rhythmicity at a lexical and grammatical level [2]. Indeed, repetition creates a certain automatization of the language [4, pg 61]. In order to prove this fact, Tannen gives the example of shadowing: “repeating what is being heard with a split-second delay” [4, pg 93]. It can sometimes occur that speakers repeat unconsciously what they have heard from another speaker, within seconds of expressing the utterance for the first time [4, pg 93]. In this way, the speaker who use repetition tries to participate in the conversation and understand the sent message.

From another perspective, Trausan-Matu showed that repetition of cue words or phrases may give birth to

artifacts that could be the key to solve problems in computer-supported collaborative learning chats [10].

### IMPLEMENTATION

A software application (Figures 1 and 2 present the graphical interface) was developed, which analyzes the rhythm in conversations. The application was tested on 10 chats in English, annotated in a specific XML format [6]. Each conversation has 4-5 participants, who are discussing in approximately 2 hours the advantages and disadvantages of the following forms of communication: blog, chat, wiki, Google Wave, and forum [6]. By analyzing the chats, it can be observed a different degree of participants' involvement, analyzed in terms of the number of utterances and words. The application was written in the Java programming language and has a graphical interface made using Java Swing and Java Applets.

For studying and comparing the rhythm in conversations, we used the following metrics:

- the total number of words;
- the total number of repeated words: each word is reduced to its stem. Two different words with the same stem will be considered a repetition (e.g. the words "user", "using" have the same stem, "us"). Function words are ignored unless they are part of a repeated phrase. An important aspect taken into account in the implementation of the repetition metric is the timestamp, i.e. the time difference between the two words repetition [4, page 64]. Thus, we have considered that two words or phrases are repeated if the distance in words between the two occurrences is less

than a constant (30 in experiment), in the case of calculating the metrics for the entire chat. At timestamps/sliding windows, where the number of lines/words is obviously smaller, the repetition of two words disregards the distance between them;

- the number of function words;
- the number of content words;
- the total number of unique function words (ignoring repetition) - calculated at the level of their stems;
- the total number of unique content words (ignoring repetition) - calculated at the level of their stems;
- the total number of unique words (ignoring repetition) - calculated at the level of their stems

The functionalities of the software application are:

1. Dividing chat in timestamps of 5 minutes (about 20 utterances). For each timestamp, the metrics defined above are computed.
2. Dividing chat in sliding windows. Dividing chat in fixed timestamps can lead to the partition of collaboration areas in which repetition plays an important role. To avoid this, every line will be the first line of a sliding window with a duration of 5 minutes. For each chat, it will be obtained a number of sliding windows approximately equal with the number of utterances. For each sliding window are computed the metrics defined above.
3. Finding the most frequent words in chat. Only content words are taken into account, function words being ignored. The stem of each word (obtained with the help of Porter Stemmer algorithm) is a key in a dictionary and its corresponding value is the number of occurrences in the

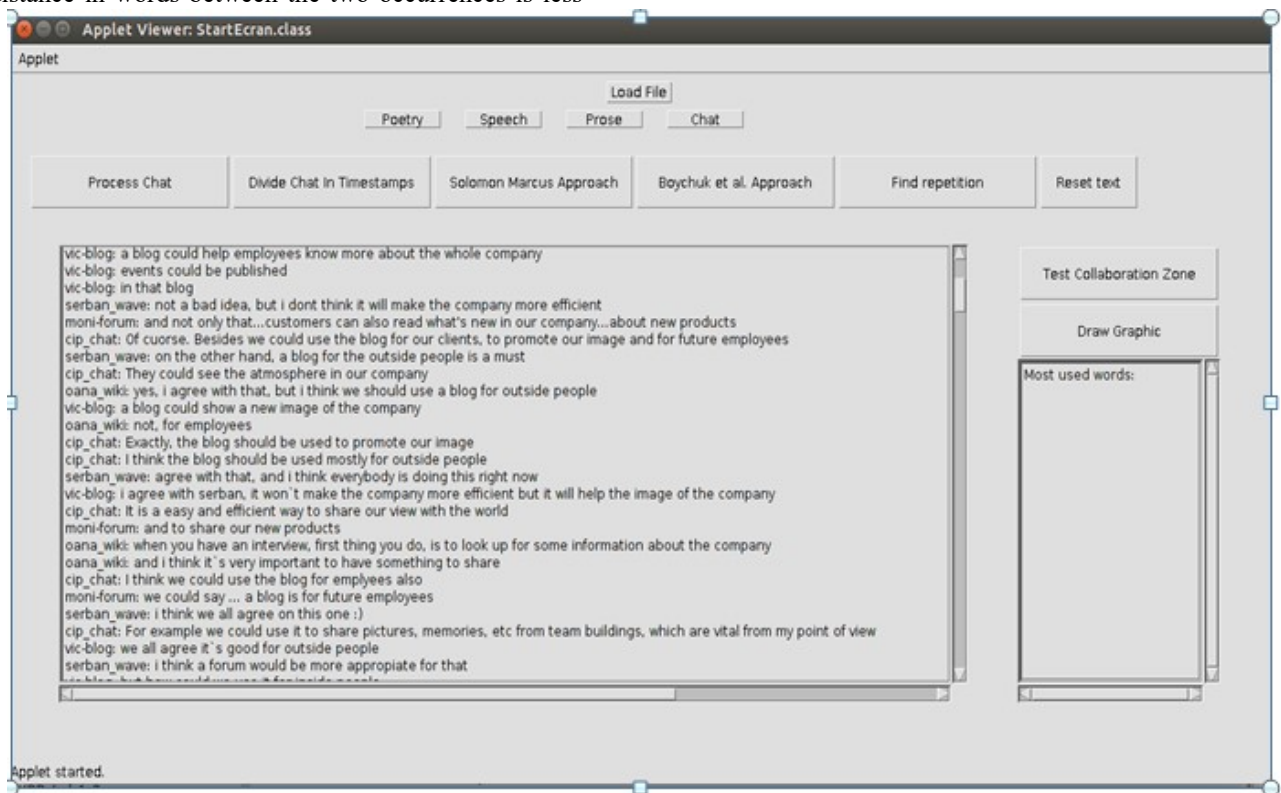


Figure 1: Graphical user interface of the application

chat.

No chat	Start ID	End ID	No. of words	No. of repeated words	Percent of repeated words
1	1	344	2832	456	16.1
2	1	296	2409	576	23.91
3	1	419	4865	795	16.34
4	1	261	3314	549	16.56
5	1	430	5533	673	12.16
6	1	392	2949	457	15.49
7	1	284	2246	431	19.18
8	1	397	3181	542	17.03
9	1	203	1428	261	18.27
10	1	311	3201	439	13.71

Table 1: Metrics calculated for the 10 chats

Tannen does between literature and conversation [4], the idea to use these two approaches in a chat appeared. In a conversation, more people exchange the role of listener and speaker, so rhythm can be analyzed at the level of the entire conversation (i.e., the utterances of all speakers, taken as a whole) and at the level of each speaker. The rhythm of a conversation can therefore be seen as a sum of the rhythm of all "voices" (in a generalized way [7, 8]) that are present in the conversation. In order to analyze the rhythm of the entire conversation, in mathematical terms, the next steps were followed:

- concatenating all the utterances from the chat to get a text ready to be processed
- dividing the obtained text in rhythmic units according to the approach chosen; in Solomon Marcus approach, the text is divided into sentences and in Boychuk et al. approach the text is delimited according to punctuation marks, coordinate conjunctions, and subordinate conjunctions.
- for each rhythmic unit, the rhythmic factors such as the rhythmic structure, the rhythmic index, the upper rhythmic limit, the lower rhythmic limit, the rhythmic diameter can be calculated using the formulas proposed by Marcus [1]

5. Applying a modified Solomon Marcus approach (at the level of words). In Solomon Marcus standard approach [1],

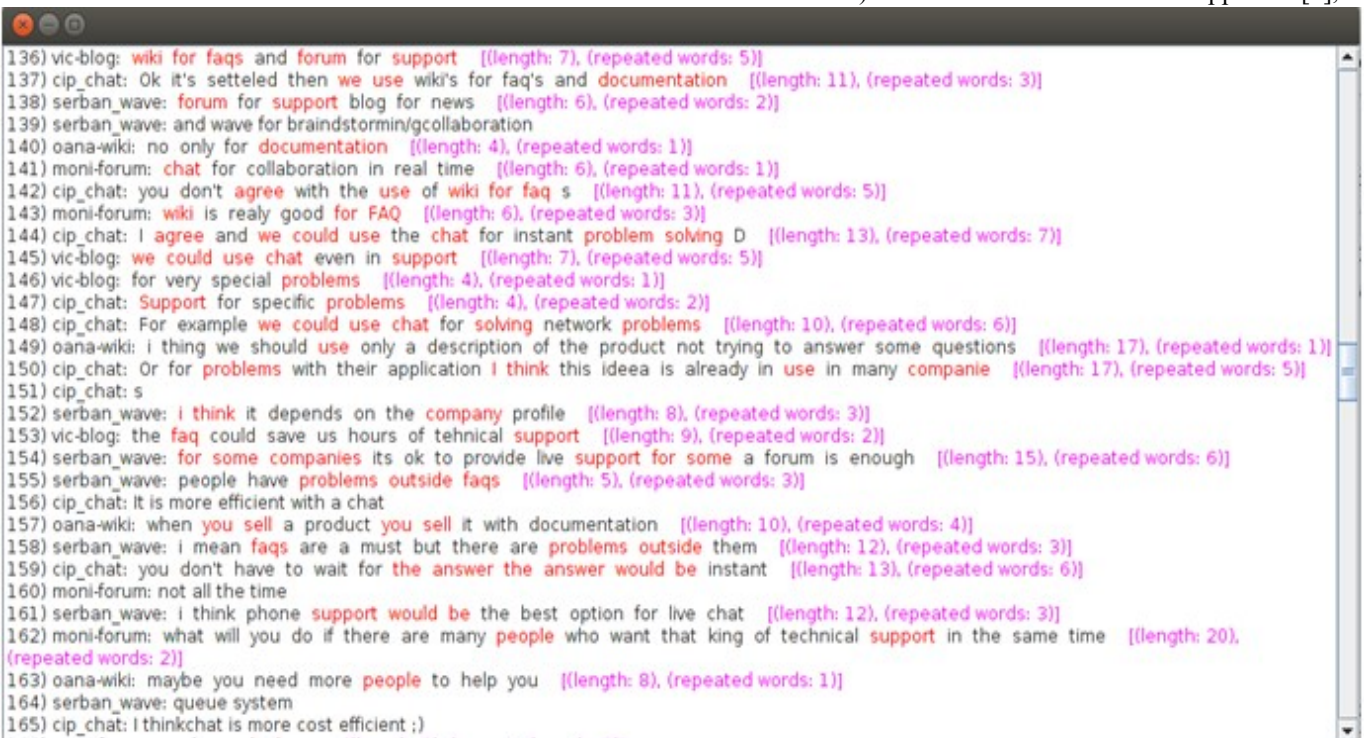


Figure 2: Highlighting repeated words/phrases in a chat

4. Applying Solomon Marcus approach [1] and Boychuck et al. approach [2]. Considering the analogy that Deborah

the rhythmic unit is divided into syllables and stressed syllables are found. The rhythmic factors are computed depending on the distance between the accents found in the

unit. To analyze the rhythm of conversation, we made an analogy between a stressed syllable and a stem frequently used in a chat. Thus, the rhythmic factors will be calculated with the same formula, for the most widely used 15 words / stems of the chat. For example, the rhythmic structure is an array, consisting of elements that represent the distance between two repeated instances of the same stem. The length of the rhythmic structure will thus be equal to the number of repetitions of the stem in the chat. A chat is seen as a single rhythmic unit (thus the length of the rhythmic unit is equal to the number of words in chat), and the stem on which this approach is applied is seen as a rhythmic nucleus.

## RESULTS

The results obtained for the ten chats are presented in Table 1. Obviously, the ratio between the number of repeated words and the total number of words is related to the value of the constant (30), which represents the maximum distance between two occurrences of a word in order to be considered as a repetition. Increasing this constant will also proportionally increase the ratio. The number of words, 30, is found in about 2-3 lines and of course, repetition can occur after a larger number of replies. For example, by setting the distance 100, the ratio becomes around 30%, and by setting the distance 200, the ratio varies around 40%. We noticed that for a distance greater than 300, the ratio does not increase significantly, remains around 40%.

It should also be noticed that the function words are ignored when counting repetitions, if they are not part of a repeated phrase such as: function word 1 + function word 2, function word + content word, content word + function word.

By dividing chats into timestamps, it was observed that the first and last utterances are introducing or ending ideas, as it is normal in a conversation, so the number of words is quite small. Both the total number of words (the maximum value obtained exceeds 400 words, but for most of the intervals, it varies between 120-250 words) and the number of repeated words increase in the following timestamps. The ratio between the number of repeated words and the number of words reaches a maximum value of 40%, but the values for most of the intervals vary between 20% - 35%. Table 2 presents the results obtained for this approach for one of the chats used for testing.

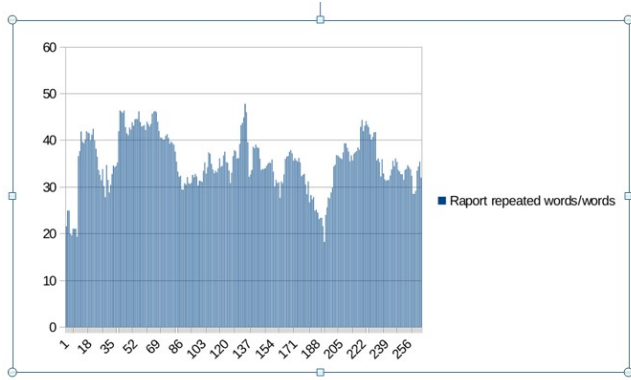
Start ID	End ID	No of words	No of repeated words
1	8	6	0
9	34	86	18
35	50	131	37
51	67	127	43
68	87	201	65
88	106	196	45
107	132	292	75
133	149	203	73
150	171	241	66
172	192	225	56
193	215	183	56
216	240	171	39
241	263	216	50
264	286	176	53
287	306	210	72
307	323	139	16
324	344	29	8

**Table 2: The results from dividing one of the chats into timestamps**

Dividing a chat in sliding windows (an example of the results obtained for this approach is presented in Figure 3) leads to the observation that the intervals or windows for which higher percentage of repetition are obtained, correspond with the collaborative areas manually annotated [5]. In this paper, the same chats are analyzed than those in a previous one, with different methods [5]. Thus, we can conclude that repetition is most common in the parts of the chat where participants discuss and support their view in a dynamic and enthusiastic way. Moreover, in conversations, participants reuse the same words, instead of synonyms, probably because they were recently heard.

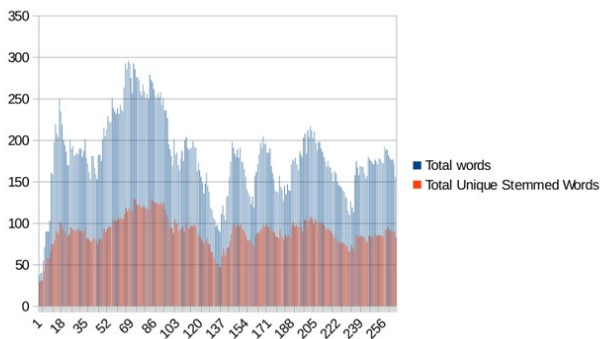
It was also noted that in all the 10 chats used for testing, the number of unique words (calculated at the level of stems) is about half the total number of words in the approach of sliding windows. Figure 4 highlights this observation.





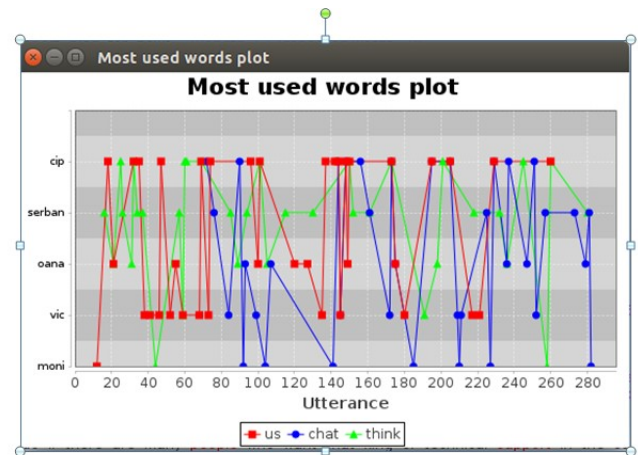
**Figure 3: The ratio between the number of repeated words and the number of words computed for one of the chats, using the approach of sliding windows**

The most repeated words in chats are: "blog", "chat", "us" (the stem for "use", "using", "user", etc.), "think", "agre" (the stem for "agree" and the words derived from it), "wiki", "ye" (the stem for "yes"), "inform", "good", "wave", "forum" etc. The presence of the words "blog", "chat", "wave", "forum", "wiki" list is explained by the fact that the chats are a discussion of these forms of information transmission and communication, so it is normal that these words are used frequently. In this case, we did not take into account the distance between repetitions. Furthermore, the repetition is given by the occurrences of the stems of the words, not by the occurrences of the words themselves. As an example, Figure 5 is a graphic that outlines the frequently used terms in one of the chats used for testing.



**Figure 4: Diagram illustrating the number of words and the number of unique words (calculated at the level of the stems) in one of the chats, using the approach of sliding windows**

After applying Solomon Marcus and Boychuck et al. approaches, we obtained the following results (Table 3 and Table 4), computed at the level of the entire chat.



**Figure 5: Graphical representation of frequently used words in one of the chats according to their occurrences in the participants' utterances**

Chat	The rhythmic index	The rhythmic upper limit	The rhythmic lower limit	The rhythmic diameter	The number of rhythmic units
1	8	14	1	13	443
2	6	14	1	13	314
3	8	13	1	12	551
4	5	13	1	12	360
5	6	12	1	11	513
6	7	15	1	13	396
7	5	11	1	9	294
8	9	16	1	15	421
9	5	16	1	14	195
10	9	16	1	15	356

**Table 3: Results for Solomon Marcus's approach, considering the whole chat as a text**

Chat	The rhythmic index	The rhythmic upper limit	The rhythmic lower limit	The rhythmic diameter	The number of rhythmic units
1	8	14	1	12	626
2	6	10	1	9	600
3	8	13	1	12	1148
4	5	12	1	11	730
5	6	12	1	10	1198
6	7	12	1	11	731
7	7	11	1	9	552
8	8	11	1	9	745
9	4	10	1	9	314
10	6	14	1	13	677

**Table 4: Results for Boychuck et al. approach, considering the whole chat as a text**

After applying the modified Solomon Marcus approach, it was observed that the rhythmic index obtained for the most used word (stem) has a value between 60-80. Obviously, as the length of the rhythmic unit is the number of words in the chat, the index increases inversely proportional with the rhythmic length. The lower rhythmic limits for the 15 most used words are less than 10, and in their corresponding rhythmic structures, elements less than 20 can be observed (a proof that repetition occurs at a low timestamp).

## CONCLUSIONS

In order to analyze rhythm in conversations, the research was based on the previously used approaches for prose, poetry and political discourse [9], as well as on the repetition of words [10]. The results obtained for the ten chats show that the areas of significant repetition correspond with collaborative areas. Moreover, it was noted that the results for Solomon Marcus's approach are similar with the results obtained for Boychuck et al approach. The

percentage of words repetition vary between 12% - 24% at the level of the entire chats, and the same percentage reaches up to 40% for the sliding windows approach.

In conclusion, natural language processing is helpful in order to analyze texts rhythm, as by automating the process of finding rhythm, comparisons between different texts can be made.

**Acknowledgments.** This research was partially supported by the FP7 2008-212578 LTfLL project.

## REFERENCES

1. Marcus, S., *Poetica Matematică*, Academy of the Socialist Republic of Romania Publishing House, Bucharest (1970).
2. Boychuk, E., Paramonov, I., Kozhemyakin, N. and Kasatkina, N., *Automated Approach for Rhythm Analysis in French*, *Proceeding of the 15th Conference of FRUCT Association* (Finnish-Russian University Cooperation in Telecommunications. (2014)
3. Dinu, M., *Ritm și rimă în poezia românească*, Cartea Românească Publishing House, Bucharest, (1986).
4. Tannen, D., *Talking Voices - Repetition, Dialogue and Imagery in Conversational Discourse*, Cambridge University Press, Second Edition (2007).
5. Dascălu, M., Trăușan-Matu, S., Dessus, P., and McNamara, D.S., *Dialogism: A Framework for CSCL and a Signature of Collaboration* (2015).
6. Trăușan-Matu, S., Dascălu, M., Rebedea, T.E., Gartner, A., *Corpus de conversații multi-participant și editor pentru adnotarea lui*, *Revista Română de Interacțiune Om-Calculator* 3(1) (2010), 53-64
7. Trausan-Matu, S., Stahl, G., Sarmiento, J., *Supporting Polyphonic Collaborative Learning*, *E-service Journal*, vol. 6, nr. 1, Indiana University Press, (2007), pp. 58-74.
8. Trausan-Matu, S., Dascalu, M., Rebedea, T., *PolyCAFe—automatic support for the polyphonic analysis of CSCL chats*, *International Journal of Computer-Supported Collaborative Learning*, 06/2014; Volume 9(2), Springer, (2014), pp. 127-156
9. Niculescu, I.D., Trausan-Matu, S., *Rhythm analysis of texts using Natural Language Processing*. In A. Iftene & J. Vanderdonck (Eds.), *Romanian Conference on Human-Computer Interaction (RoCHI 2016)*, (2016), pp.197-112
10. Trausan-Matu, S., *Repetition as Artifact Generation in Polyphonic CSCL Chats*, *Third International Conference on Emerging Intelligent Data and Web Technologies*, IEEE Conference Publications, pp. 194-198 (2012)

# Using Text Processing in a Multimedia Environment. UAIC Activities in MUCKE Project

Adrian Iftene

Faculty of Computer Science,  
 “Alexandru Ion Cuza” University of Iasi  
 General Berthelot, No. 16  
 adiftene@info.uaic.ro

## ABSTRACT

This paper presents main activities of UAIC (“Alexandru Ioan Cuza” University) team from the MUCKE project. MUCKE addressed the stream of multimedia social data with new and reliable knowledge extraction models designed for multilingual and multimodal data shared on social networks. Credibility models for multimedia streams are a novel topic, and constituted the main scientific contribution of the project. UAIC group was involved in the main tasks of the project: building the data collection, text processing, diversification in image retrieval and data credibility.

## Author Keywords

Image retrieval; text processing; diversification; YAGO; Flickr

## ACM Classification Keywords

H.5.2. Information interfaces and presentation (e.g., HCI): User Interfaces. H.3.2. Information Storage and Retrieval: Information Storage.

## General Terms

Human Factors; Design.

## INTRODUCTION

MUCKE (Multimedia and User Credibility Knowledge Extraction) project of type ERA-NET CHIST-ERA started in 2012 and finished at the end of 2015. MUCKE departed from current knowledge extraction models, which are mainly quantitative, by giving a high importance to the quality of the processed data, in order to protect the user from an avalanche of equally topically relevant data. MUCKE came with two central innovations: automatic user credibility estimation for multimedia streams and adaptive multimedia concept similarity. Adaptive multimedia concept similarity departed from existing models by creating a semantic representation of the underlying corpora and assigning a probabilistic framework to them.

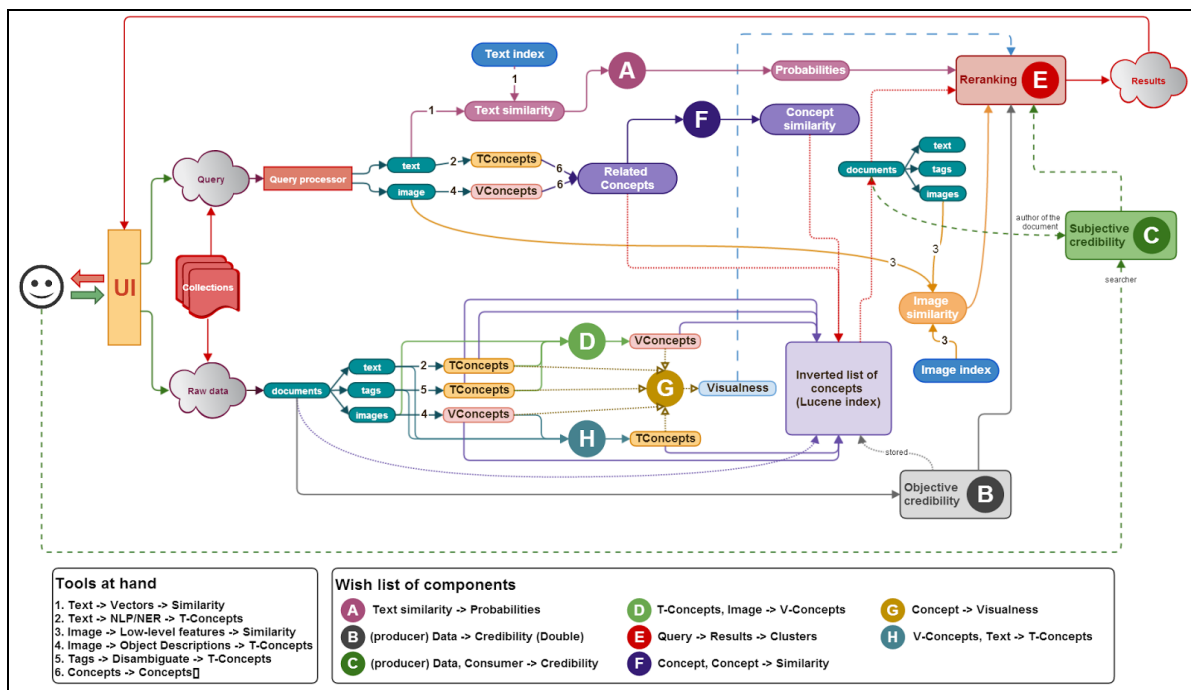


Figure 1. MUCKE Architecture [1].

Although a lot of research efforts were directed towards automatic information mining and important results are reported in different domains, the topic remains complicated with a huge potential. Important challenges arise from the heterogeneous character of raw data, from the scalability of processing methods and from the reliability of extracted knowledge.

Heterogeneity comes in different forms: nature of the documents (text, image, speech, movie), diversity of the languages used in text documents, or language particularities for different data sources. Scalability remains an the most important issue for multimedia streams, whose processing with a good quality and exploitation under real time constraints are still problematic. While it is difficult for the results of automatic techniques to match the quality of manually created resources, the latter imply huge investments when dealing with large-scale data.

Figure 1 shows an overview of the MUCKE framework, covering how documents are processed, concepts extracted and indexed, similarity computed based on concepts, text and images, and how credibility is estimated and fed into the re-ranking process to and how credibility is estimated and fed into the re-ranking process to improve the final set of results. In the rest of the paper, we will see the main tasks in which UAIC group was involved.

### NEW DATA COLLECTION

The complex nature of the project objectives required the mobilization of different multimedia data sources in order to mine all necessary user-related knowledge. With over 6 billion of photo uploads, Flickr (flickr.com) is one of the largest photo repositories on the Web and constituted our main source of visual information. Consortium members downloaded images and textual metadata from 1,000,000 Flickr users.

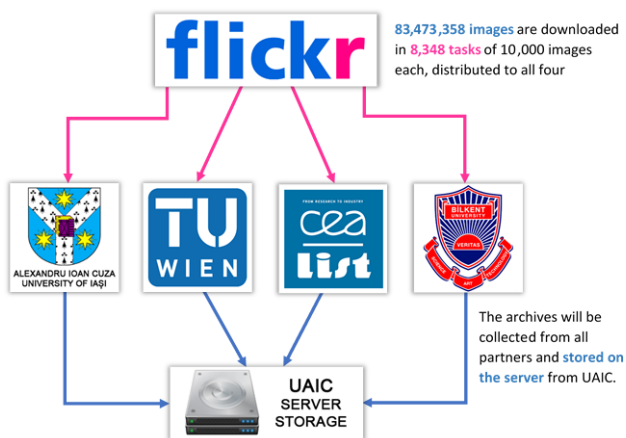


Figure 2. How the download process works.

Given that Flickr contains a mix of personal and social relevant data, focus was on downloading the latter type, which is useful for information extraction tasks. UAIC

coordinated the data collection effort but, in order to speed up the process, it was distributed among all partners (Figure 2).

During the download process, some statistics were provided in the form of two charts: one for ongoing and one for completed tasks. These charts were dynamically updated, thus enabling us to check our tasks and see the overall status of the download process at any time (See Figure 3).

At the end of downloading process, the MUCKE corpus contained metadata and images based on Wikipedia concepts that are often used to annotated Flickr images. Wikipedia concepts are ranked using the number of corresponding Flickr images which is divided by the log of incoming Wikipedia links in order to penalize very common concepts. In the end, the top 200 concepts with the highest frequency of occurrence in Flickr are represented in the current corpus.

At the end of MUCKE project, we collected over 80 million images and their associated metadata that have been downloaded mainly from the Flickr database.

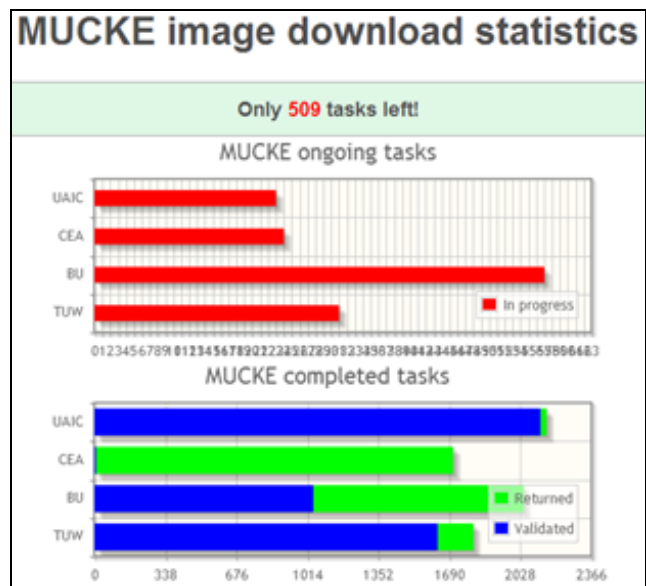


Figure 3. Image download statistics during the download process.

### BASIC TECHNIQUES FOR TEXT PROCESSING

Before any keywords or query semantic analysis is performed, some shallow linguistic processing steps are commonly required, such as lemmatization, Part-of-Speech (POS) tagging, anaphora resolution, named entity identification. The aim of text processing was to help us to map images metadata to Wikipedia concepts.

#### Lemmatization

Lemmatization identifies the root word beyond the inflected forms, which is necessary since most concepts from Wikipedia contain root words. This allows us access to both linguistic resources such as WordNet [28] (for word senses

and semantic relations) and to other linguistic processing tools using root word lexicons or patterns. For all shallow processing steps we identified free existing tools, for English, French, German and Romanian [16, 17].

### POS-Tagging

POS-tagging is another important pre-processing step, identifying the parts of speech of the words in the target document. Usually POS-taggers also add morpho-syntactic annotations to the words. This processing step is essential for any deeper analysis, since it conveys some data regarding the grammatical relations between the words, usually correlated to semantic relations. For lemmatization and POS-tagging we used a web-service [25] found to perform best for the Romanian language [26] used in UAIC's experiments, but all European languages have equivalent quality tools available.

### Identifying Noun-Phrases and Named Entity

Identifying Noun-Phrases and Named Entity is another step relevant for our project, because it helps us to identify Wikipedia articles (in fact Wikipedia concepts). For Romanian, we used tools from [8, 9, 14 and 18].

Additionally, for named entity recognition we used web service described in [25], with two rules that allow us to transform proper names into named entities. The rules were these: (1) many successive capitalized words and marked by POS-Tagger as proper names are grouped into a single entity, (2) many capitalized words separated by linking words such as "din", "de", etc. (in En: from, of) were also grouped into one entity (for example "Venus din Milo" (in En: Venus from Milo), "Camera Națională de Pensii Publice" (in En: National Chamber of Public Pensions), etc.). We validated these entities using Wikipedia or our external resources (an entity is considered valid if it has a corresponding Wikipedia page or it exists in our external resources).

For named entities classification, we looked at the words in the neighborhood of the entity in order to make the classification. Thus, in the situations where we found in the text expressions such as "Palatul Roznovaru" (in En: Roznovaru Palace) or "munții Rodnei" (in En: the Rodnei Mountains), we considered the corresponding type "palace" or "mountain" for them. In the other situations, we used our external resources, where we have entities of the type location, organization, people and other.

### Anaphora resolution

Anaphora Resolution identifies semantic identity between different parts (usually NPs) of a text. We did the anaphora resolution both in the original text and in the Wikipedia articles [12]:

- (1) In the original text we started from the classification of the named entity. For instance, after we classified "Munții Rodnei" as "munte", all

the appearances of the word "munte" (En: mountain) (that were not followed by the word "Rodnei") have been replaced by the expression "Munții Rodnei".

- (2) At the level of Wikipedia articles, we consider that the first paragraph refers to the presented concept from the current Wikipedia page. Thus, we considered that all expressions such as: "A fost inaugurat" (in En: was inaugurated), "Este conceput" (in En: is designed/conceived), "Este considerat" (in En: is considered), etc. refer to the main concept, described in that Wikipedia page.

## ADVANCED TECHNIQUES USED FOR TEXT PROCESSING

### Query reformulation

Query reformulation provides techniques of improving the quality of search results by extending or replacing parts of the original query. This process is very similar to a required step in a question answering system as described in [15].

Here, we apply two approaches:

- (1) a global technique, which analyses the body of the query in order to discover word relationships (synonyms, homonyms or other morphological forms from WordNet), to remove stop words ("a", "un", "la", "pentru", (English: the, a, at, for), etc.), to remove wh- words ("cine", "ce", "de ce", "unde", (English: who, what, why, where), etc.) and to correct any spelling errors;
- (2) local feedback which implies the analysis of the results returned by the initial query, leading to re-weighting the terms of the query and relating it with entities and relationships originating from the target ontology. Further discussion on expanding search queries using Yago and Wikipedia can be found in [13].

Another experiment carried out at UAIC with the goal of augmenting the existing query using the information extracted from large data resources such as Wikipedia and Freebase [5]. The system uses a POSTagger for the Romanian language, afterwards it identifies and classifies the named entities. For these entities, the external resources mentioned above are used and concepts are identified (an entity that can appear in the text in various forms) and the relations between them. Further details about this experiment can be found in [7].

### Diversification with YAGO

From all collected data for MUCKE, we selected around 30.000 images (a small collection), with aim to perform several processing tasks at both textual (on associated metadata) and image level and retrieve the results in a diversified way. Over small collection, we built a system which allowed users to retrieve multimedia content [11].



To improve a search query we first looked for the relevant words in the query in the results provided by a text-processing module. The text processing module is used to process on one hand, the images associated metadata and, on the other hand, the user queries. Standard tools are used for POS-tagging [25], lemma identification [26] and named entity identification [9]. After the images associated metadata is processed, the image collection is indexed with Lucene [21]. In order to achieve diversification in the results set, the system incorporates a query expansion module that makes use of the YAGO ontology [29] (see Figure 4).

Yago ontology comprises well known knowledge about the world [10]. It contains information extracted from Wikipedia [27] and other sources like WordNet [28] and GeoNames [6] and it is structured in elements called entities (persons, cities, etc.) and facts about these entities (which person worked in which domain, etc.).

For example, with Yago we are able to replace in a query like “tennis player on court”, where we have two entities (“tennis player” and “court”), the entity “tennis player” with instances like “Roger Federer”, “Rafael Nadal”, “Andy Murray”, etc. Thus, instead of performing a single search with the initial query, we perform several searches with the new queries, and in the end we combine the obtained partial results in a final result set.

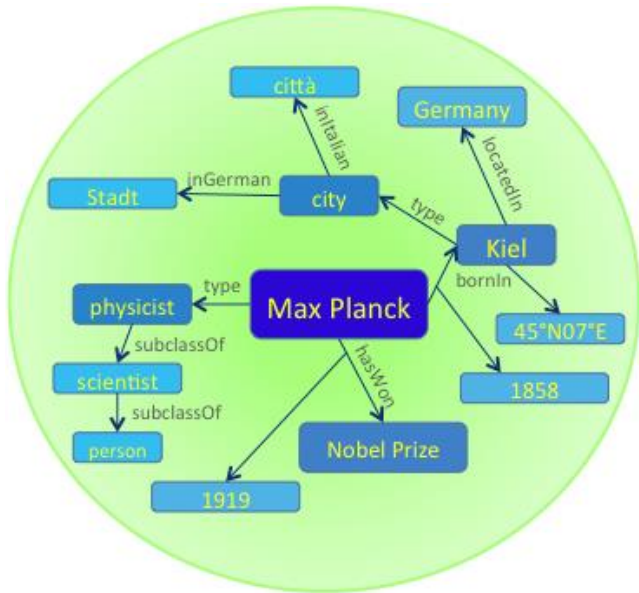


Figure 4. YAGO Ontology [29].

In Figures 5 and 6 are presented results obtained with “tennis player on court” query in our application and in Google.

How we can see, results offered by our system contains both concepts (“tennis player” and “court”), while in the Google are cases when concept “tennis player” is missing.

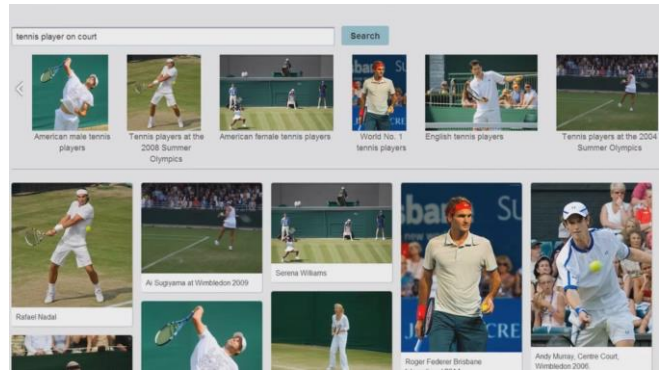


Figure 5. Results for query “tennis player on court” in our application.

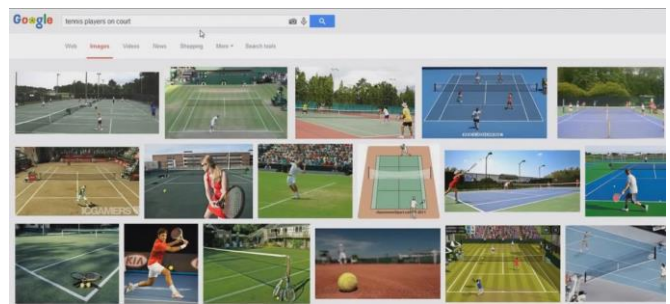


Figure 6. Results for query “tennis player on court” in Google application.

### AUTOMATIC IMAGE ANNOTATION

The aim of this component is to add relevant keywords to an image without annotation [18]. In order to do this, the first step was to create a collection of images that was annotated by human annotators, while the second step was to expand this collection of images performing search on the Internet using keywords associated to the initial collection of annotated images. Currently, for a new picture, we can identify similar images in our collection of images and based on the keywords associated with them, we can determine what keywords characterize this new image.

### Creation of gold collection with annotated images

The initial collection of images consisted of 100 images, from different areas (art, furniture, sport, other, etc.). The images were selected by six human experts and then were manually annotated by human annotators. Some of the images have words in their visual content to see how this can influence the process of annotation. In the process of annotating, the only criterion was to write keywords in the Romanian language, criterion that was established from the beginning.

Comparing the keywords entered by users for the same picture, it was seen that there were small differences among the words entered, most of them were from the same lexical

family or they were synonyms. Performing an analysis on what users have annotated, it can be said that their tendency was to introduce, on average, 3.41 keywords per image, with a minimum of 2 keywords for an image and a maximum of 12 keywords for an image. Looking further into the keywords that they have entered, it can be said that most users have opted for simple words and not phrases.

After completing this step, we increased the initial collection as it follows: for each image we added 10 new images to our collection, thus increasing the image collection to 1,000 images. For this, we searched for Google images using lists of keywords associated with each image. For the first 10 results, we initially associated the list of keywords used in the search process, followed by a process of verification, corrections, additions to this list, this process was done with human annotators [19].

### Reverse Image Search

This module uses the 1,000 collection of images with related keyword lists obtained at the previous step. Regarding this collection, we know that the list contains relevant keywords associated with images.

The main purpose of this module is to generate a list of keywords that characterize an image given by the user. This is done using the LIRE library [20] (Lucene Image REtrieval) [22], which compares the new image with the images from our collection. It establishes a set of 20 images most closely to the new inserted image in terms of texture and color. After that, starting from the 20 lists of keywords associated to these images, we use lemmatization, the synonymy relation and the processing of expressions. In the end, using keywords order in the lists and their frequency, we decide the list keywords for new image.

The conclusion that can be drawn from the analysis of obtained results: the more images we have in our collection of annotated images, the more chances of finding similar images.

### CONCLUSIONS

In present, it is a growing interest in hosting multimedia information and looking for such content. Large companies like Flickr, Google, Bing, Yahoo, and Microsoft have such platforms available to users both through its own search pages, and through APIs available to application developers. In order to create collection of images we needed in the project, we used the Flickr network (for images) and Wikipedia encyclopedia (to establish concepts). For Flickr API, we created a crawler which allowed us to save locally relevant information: both images and associated metadata. Since downloading the massive amount of information required a huge effort on our part, this activity was implemented with all partners involved in the project MUCKE.

We could also see that text processing techniques and related resources can make a significant contribution to

improving quality of the image retrieval task. The principal methods of doing this come either from improving the way it is used query (extension of search terms, finding semantic relations between terms in the query and use clues from the image, use multilingualism to expand the scope of search) or on improving description of an image. For this, we have identified and used several tools for text processing, such as POS taggers, tools for lemmatization, for anaphora resolution, for identification of name entities, etc.). We also used semantic resources such as WordNet, Yago, Wikipedia, GeoNames and Freebase. The quantitative evaluation was performed for all experiments described with good results, more details are available in [7, 13 and 18].

Our algorithms, presented in this paper, have been successfully used in the Plant Identification task from CLEF [3], in Image CLEF evaluation campaigns (in Scalable Concept Image Annotation Challenge task and Plant Identification task) and in MediaEval benchmark [2, 4, 9, 12, 23 and 24].

### ACKNOWLEDGMENT

We want to thank to UAIC team, to bachelor and master students from Faculty of Computer Science, “Alexandru Ioan Cuza” University of Iasi, which was involve in the all stages of this project development. This work is partially supported by POC-A1-A1.2.3-G-2015 program, as part of the PrivateSky project (P\_40\_371/13/01.09.2016).

### REFERENCES

1. Bierig, R., Șerban, C., Siriteanu, A., Lupu, M, and Hanbury, A. A System Framework for Concept- and Credibility-Based Multimedia Retrieval. In *ICMR 2014 Proceedings of International Conference on Multimedia Retrieval*, Glasgow, Scotland, (2014), 543-550.
2. Calfa, A., Silion, D., Bursuc, A. C., Acatrinei, C. P., Lupu, R. I., Cozma, A. E., Pădurariu, C. and Iftene, A. Using Textual and Visual Processing in Scalable Concept Image Annotation Challenge. In *Working Notes of CLEF 2015 - Conference and Labs of the Evaluation forum - ImageCLEF2015*, 1391, 8-11 September 2015, Toulouse, France, (2015).
3. CLEF 2017: <http://clef2017.clef-initiative.eu/> (Last time accessed on 30 June, 2017)
4. Cristea, A.G., Savoia, M. M., Martac, M. A., Pătraș, I. C., Scutaru, A. O., Covrig, C. E. and Iftene, A. Using Machine Learning Techniques, Textual and Visual Processing in Scalable Concept Image Annotation Challenge. In *Working Notes of CLEF 2016 - Conference and Labs of the Evaluation forum - ImageCLEF2016*, Evora, Portugal, (2016).
5. Freebase: <https://www.freebase.com> (Last time accessed on 30 June, 2017)
6. GeoNames: <http://www.geonames.org/> (Last time accessed on 30 June, 2017)

7. Gherasim, L. M. and Iftene, A. Extracting Background Knowledge about World from Text. In *Proceedings of the 10th International Conference "Linguistic Resources and Tools for Processing the Romanian Language"*, Craiova, 18-19 September 2014, ISSN 1843-911X, (2014), 199-208.
8. Gînscă, A. L., Boroş, E., Iftene, A., Trandabăţ, D., Toader, M., Corîci, M., Perez, C. A. and Cristea, D. Sentimatrix - Multilingual Sentiment Analysis Service. In *Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis (ACL-WASSA2011)*, Portland, Oregon, USA, (2011).
9. Gînscă, A. L., Popescu, A., Lupu, M., Iftene, A. and Kanellos, I. Evaluating User Image Tagging Credibility. *Experimental IR meets Multilinguality, Multimodality, and Interaction*, Lecture Notes in Computer Science, 9283, Publisher Springer International Publishing. In *Proceedings of 6th International Conference of the CLEF Association, CLEF'15 Toulouse, France*, September 8-11, (2015), 41-52.
10. Hoffart, J., Suchanek, F., Berberich, K. and Weikum, G. YAGO2: A spatially and temporally enhanced knowledge base from Wikipedia. *Elsevier, Artificial Intelligence*, 194, (2013), 28-61.
11. Iftene, A. and Alboaic, L. Diversification in an image retrieval system based on text and image processing. *Computer Science Journal of Moldova*, 22 (3) (66), (2014), 1-10.
12. Iftene, A., Moruz, A. and Ignat, E. Using Anaphora resolution in a Question Answering system for Machine Reading Evaluation. *Notebook Paper for the CLEF 2013 LABs Workshop - QA4MRE*, 23-26 September, Valencia, Spain, (2013).
13. Iftene, A., Siriţeanu, A. and Petic, M. How to Do Diversification in an Image Retrieval System. In *Proceedings of the 10th International Conference Linguistic Resources and Tools for Processing the Romanian Language*, Craiova, 18-19 September 2014, (2014), 153-162.
14. Iftene, A., Trandabăţ, D. and Pistol, I. Grammar-based automatic extraction of definitions and applications for Romanian. In *Proceedings of RANLP workshop "Natural Language Processing and Knowledge Representation for eLearning environments"*, (2007), 978-954.
15. Iftene, A., Trandabăţ, D., Moruz, A., Pistol, I., Husarciuc, M. and Cristea, D. Question answering on English and Romanian languages. In *Workshop of the Cross-Language Evaluation Forum for European Languages*, Springer Berlin Heidelberg, (2009), 229-236.
16. Iftene, A., Trandabăţ, D., Moruz, A.-M., Pistol, I., Husarciuc, M. and Cristea, D. Question Answering on English and Romanian Languages. In *Multilingual Information Access Evaluation, Vol. I Text Retrieval Experiments*, Springer, Heidelberg, CLEF 2009, LNCS 6241, Part I, (2010), 229-236.
17. Iftene, A., Trandabăţ, D., Pistol, I., Moruz, A. M., Husarciuc, M. and Cristea, D. UAIC participation at QA@ CLEF2008. In *Workshop of the Cross-Language Evaluation Forum for European Languages*, Springer Berlin Heidelberg, (2008), 385-392.
18. Laic, A. and Iftene, A. Automatic Image Annotation. In *Proceedings of the 10th International Conference "Linguistic Resources and Tools for Processing the Romanian Language"*, Craiova, 18-19 September 2014, ISSN 1843-911X, (2014), 143-152.
19. Laic, A., Gherasim, L. M. and Iftene, A. Expanding a gold collection of images using the Flickr network. In *Proceedings of the Conference on Mathematical Foundations of Informatics MFOI'2016*, July 25-29, 2016, Chisinau, Republic of Moldova, (2016), 241-250.
20. LIRE: <http://www.semanticmetadata.net/lire/> (Last time accessed on 30 June, 2017)
21. Lucene: <http://lucene.apache.org/> (Last time accessed on 30 June, 2017)
22. Lux, M. and Marques, O. Visual Information Retrieval using Java and LIRE. *Synthesis Lectures on Information Concepts, Retrieval, and S*, Morgan & Claypool Publishers, 112 pages, (2013).
23. MediaEval Benchmark: <http://www.multimediaeval.org/>
24. Şerban, C., Siriţeanu, A., Gheorghiu, C., Iftene, A., Alboaic, L. and Breabăn, M. Combining image retrieval, metadata processing and naive Bayes classification at Plant Identification 2013. *Notebook Paper for the CLEF 2013 LABs Workshop - ImageCLEF - Plant Identification*, 23-26 September, Valencia, Spain, (2013).
25. Simionescu, R. Hybrid POS Tagger. In *Proceedings of Language Resources and Tools with Industrial Applications*, Workshop (Eurolan 2011 summerschool), (2011).
26. UAIC NLP Tools: <http://nlptools.infoiasi.ro/WebPosRo/> (Last time accessed on 30 June, 2017)
27. Wikipedia: [http://en.wikipedia.org/wiki/Main\\_Page](http://en.wikipedia.org/wiki/Main_Page) (Last time accessed on 30 June, 2017)
28. WordNet: <http://wordnet.princeton.edu/> (Last time accessed on 30 June, 2017)
29. Yago ontology: <https://www.mpi-inf.mpg.de/departments/databases-and-informationssystem/research/yago-naga/yago/> (Last time accessed on 30 June, 2017)

# Perceived usefulness of Facebook for university students: A gender analysis across two countries

Costin Pribeanu<sup>1,2</sup>, Alexandru Balog<sup>1</sup>

<sup>1</sup> ICI Bucharest, Romania

Bd. Mareașal Averescu nr.8-10, Bucharest

<sup>2</sup> Academy of Romanian Scientists

Splaiul Independentei nr.54, Bucharest  
pribeanu@ici.ro, alexb@ici.ro

Vincentas Lamanauskas<sup>3</sup>, Violeta Slekiene<sup>3</sup>

<sup>3</sup> Siauliai University

P.Visinskio Street 25, Siauliai, Lithuania  
v.lamanauskas@esf.su.lt, violeta@fm.su.lt

## ABSTRACT

The usefulness of Facebook for the university students has been widely researched in the last decade. Existing studies reveal a diversity of potential educational benefits, such as better student engagement, better student integration in the community, and valuable support for social learning. Few studies exist that analyze the gender differences and cultural differences in using Facebook. This paper aims at analyzing the gender differences as regards the Facebook usefulness perceived by university students. The perceived usefulness has been conceptualized as a global factor with three dimensions: social, information, and collaboration usefulness. An invariance analysis has been carried on to assess the invariance of the measurement model across gender using samples from two countries: Romania and Lithuania. In both countries, students perceived Facebook as more useful for collaboration. The results show that in Romania female students have a higher perception of Facebook usefulness. A third finding is that Facebook has been perceived as being more useful by the Romanian students than by the Lithuanian students.

## Key words

Facebook, social networking websites, gender differences, cultural differences.

## INTRODUCTION

The popularity of Facebook among university students stimulated the research on the educational potential of social networking websites. Several studies have been published in the last decade that highlight the potential educational benefits, such as better student engagement [13, 14], better student integration in the community [21], useful information sources [15, 16], and valuable support for social learning [5].

Few studies exist that analyze the gender differences and even fewer (if any) that analyze the cultural differences as regards Facebook use by university students.

A shortcoming of many studies that are reporting gender differences by applying traditional statistical significance tests is the assumption that respondents are interpreting the variables in the same way. This approach may work with independent variables but may lead to ambiguous results if the variables under consideration are measures of an underlying model. In this case, an invariance analysis is needed [6, 24, 27].

The purpose of this paper is to analyze the gender and cultural differences as regards the usefulness of Facebook. The study is focusing on the usefulness for educational purposes. In this respect, it does not include the usefulness for socialization or entertainment. The perceived usefulness has been conceptualized as a global factor with three dimensions: social, information, and collaboration usefulness. In order to test the model on samples from different countries, an evaluation instrument has been developed in the context of a cooperation between researchers from ICI Bucharest (Romania) and Siauliai University (Lithuania).

The paper has two main objectives. The first is to analyze the gender differences in each country. The second objective is to analyze if there are cultural (cross-national) differences for each gender. Two data samples have been collected: one from Romania (N=758) and the second from Lithuania (N=297). Then, an invariance analysis across gender has been carried on for each sample.

The rest of this paper is organized as follows. The following section presents related work with a focus on gender differences and Facebook usefulness. In the next two sections, the method and results are presented. The paper ends with a conclusion and future work.

## RELATED WORK

Social media technologies are reshaping the way students are learning and interacting with each other. Facebook is the most popular social networking website among university students and is a challenge for educators to find the most suitable ways to exploit the educational potential.

The paradigm shift towards social learning requires considering a diversity of activities, such as meeting, active participation, critical thinking, information and content sharing [5]. As shown by Kumar [14], undergraduate students value the use of new technologies in e-learning and perceive Web 2.0 as a useful support for online discussions and engagement with their teachers and colleagues. Also, Arteaga-Sanchez et al. [3] suggest that Facebook should support collaborative and cooperative learning and may enrich the learning experience.

There are few studies that report gender differences in using Facebook in university contexts. This issue is a concern from both the computer-supported learning and e-learning aids in universities, such as the social media technologies.



Recent work in this area suggests that women are disadvantaged as regards the computer-supported learning and e-learning aids because of the inferior levels of access, technology literacy, and the dominant male behavior [11]. However, the assumption that difference implies disadvantage is challenged by the evidence that various factors, such as professed confidence and apparently dominant interaction styles, do not necessarily lead to better educational opportunity and performance. Furthermore, women often have better results than men despite the observable differences in the interaction style [11].

Evidence demonstrates that theoretical conceptions of computer-supported learning environments as offering equal opportunities are flawed because social and educational interaction that takes place through electronic channels loses none of the socio-cultural complexity or gender imbalance that exists in more traditional learning environments [28]. Such imbalances can be determined by cultural differences [10]. The impact of culture on conceptions of gender and behavior is identified as a strong determining factor of both the existence and the nature of differences [25]. So unequivocal generalization about gender differences in the usage of e-learning aids and computer-supported learning is impossible.

According to Gunn [11], the assumption that gender-based inequity as regards the access and computer literacy is a disappearing problem. This is based on the findings of his study and other various sources across the western world. The discussion turns to the differences that persist in this "more equitable" environment and possible interpretations of their impact. Losh [18] found that gender parity does exist in computer and Internet usage in the United States.

Link and Marz [17] identified gender differences in many computer-related aspects. They found that the men are using the computers more frequently, and have access to a better computer infrastructure, including the Internet connection.

Enoch and Soker [9] point to the presence of a digital divide among male-female university students in Israel, whereby males who form the dominant group are able to benefit more than the female subordinate group as the differences persist over time.

With regard to gender differences, Selwyn [23] data found female students to be significantly more likely to make use of the Internet for academic information seeking than their male counterparts. This finding runs counter to much of the research literature during the 1980s and 1990s which highlighted men's dominance in the area of education technology (for example, Sutton, [26]).

In terms of how academic information searching was patterned, female students were significantly more likely than male students to report looking for information about university studies/assignments [23]. The finding is also consistent with more general studies on Facebook usage showing that females and young people spend more time on Facebook and have more Facebook friends [20].

Donlan [8] has explored students' views on the use of Facebook in an academic context, finding, as with several previous studies, that the picture is far from clear and there

is still a considerable amount of diversity among students in terms of existing adoption and willingness to use Facebook for teaching and learning purposes. Survey findings indicated an interest among students in using Facebook for a range of teaching and learning purposes, including accessing and posting links, discussing work with others and tutorials with lecturers. There was no significant difference between male and female students in their likelihood of using social networking sites.

While exploring the use of social networking tools, Rodriguez-Hoyos et al. [22] discussed the need to widen the research by including include other dimensions such as geographical and gender differences that could affect attitudes, resistance and actual uses of these sites.

Mazman & Usluel [19] carried out a study in Turkey (N=870) and found that females are using Facebook for maintaining existing relationship, academic usage, and the following agenda more than males do.

## METHOD

### Variables and data samples

In this study, a multidimensional model of the Facebook usefulness has been used that has three dimensions: social usefulness (US), information usefulness (UI), and collaboration usefulness (UC). The conceptual model has been developed and validated in a previous work [4] on a sample collected from a Romanian university. The constructs and items are presented in Table 1.

*Table 1. Facebook usefulness for university students*

Item	Statement
US1	Using Facebook improves participation in collective activities
US2	Using Facebook I can better present my university work to other people
UI1	Using Facebook I am better informed about events of interest in my university
UI2	Using Facebook I get useful information from university people
UI3	On Facebook, I can find useful resources for my university work
UC1	Using Facebook improves communication between colleagues
UC2	Using Facebook encourages the creation of academic groups based on similar interests and needs
UC3	Using Facebook improves student group work

The questionnaire has been administered to university students from Romania and Lithuania. Students were invited to answer general questions (university, faculty, program of study, year of study, and demographics), questions regarding the Facebook use (network size, frequency, and duration of use), and then to evaluate the items on 7-points Likert scale.

The initial samples included 796 students from Romania and 303 students from Lithuania. After checking the completeness of the answers and the normality of the data, including univariate and multivariate outliers, several observations were eliminated so the working sample has 758 observations from Romania and 297 from Lithuania.

The demographics, the mean number of Facebook friends, and mean time spent daily are given in Table 2.



Table 2. Data samples

Variable	Romania (N=758)	Lithuania (N=297)
Gender M/F	418 / 340	118 / 179
Mean age	21.6 (SD=2.8)	22.3 (SD=5.3)
FB friends	756.0 (SD=816.2)	267.9 (222.2)
Min/Day	78.3 (SD=74.9)	94.3 (SD=103.8)

As shown in Table 3, in both countries female students have larger Facebook networks and spend more time daily on Facebook.

Table 3. Gender differences in usage

Variable	Romania	Lithuania
FB friends - M	684.2 (SD=737.5)	223.4 (227.2)
FB friends - F	844.3 (SD=896.9)	297.4 (227.4)
Min/Day - M	72.6 (SD=68.8)	89.7 (SD=76.3)
Min/Day - F	85.9 (SD=81.3)	97.3 (SD=118.6)

A one-way ANOVA showed that the gender differences are statistically significant in both countries as regards the number of Facebook friends and only in Romania as regards the time spent daily.

**Methodological approach**

The measurement model specification is presented in Figure 1.

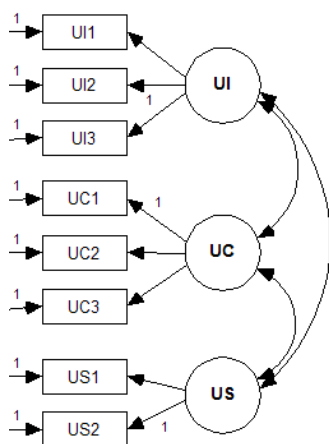


Figure 1. The measurement model

The unidimensionality has been assessed by examining the loadings ( $\lambda$ ) and t-values. The scale reliability has been assessed with the Cronbach's alpha. The convergent validity has been analyzed by examining the composite reliability (CR) and average variance extracted (AVE), following the recommendations in [1, 12].

In order to analyze the gender differences across the two countries, a methodological approach is required that is based on invariance analysis. The invariance analysis checks if the constructs are equivalent across groups [6, 27]. Lack of invariance across groups makes not possible to assess if the differences are due to different perceptions or to different interpretations of the evaluation instrument. In other words, the conclusions based on the measurement scale are ambiguous if not erroneous [27].

A multi-group CFA (MGCFA) using AMOS for Windows [2] has been conducted to check if the scale is invariant across gender. MGCFA is based on testing a hierarchical

series of nested models, starting with a baseline model that fits all the samples together. The parameters are freely estimated and a baseline chi-square value is derived.

The nested models are obtained by adding constraints for invariance [6, 27]. Two tests have been used:  $\Delta\chi^2$  (invariance if nonsignificant) and  $\Delta CFI$  (invariance if the depreciation is less than 0.01) [7].

Metric invariance tests if the factor loadings are equivalent across groups (equality of scaling units). Metric invariance enables the comparison of the observed scores.

The scalar invariance tests the equality of the intercepts (origin of the scale) across groups. Scalar invariance enables the comparison of latent mean scores. The next test adds constraints for the equivalence of the structural covariance of factors (dimensions) and, if successful, enable the comparison of the inter-factor correlations.

Prior to carrying on the multi-group CFA, the model has been tested and validated on each group, in order to check the configural invariance. Configural invariance means that in each group the dimensions are perceived in a similar way (same pattern of free and fixed factor loadings on the items) [24, 27]).

**RESULTS**

**Romanian sample**

The first step is to test the model for each gender group. The descriptive statistics for the sample and each gender is presented in Table 4.

Table 4. Descriptive statistics (N=758)

Item	Male (418)		Female (340)		Total	
	M	SD	M	SD	M	SD
US1	4.47	1.57	4.68	1.50	4.56	1.54
US2	3.93	1.59	3.97	1.59	3.95	1.59
UI1	4.68	1.54	4.96	1.52	4.81	1.54
UI2	4.78	1.44	4.79	1.51	4.79	1.47
UI3	4.43	1.68	4.63	1.60	4.52	1.65
UC1	4.99	1.65	5.07	1.64	5.02	1.65
UC2	4.92	1.59	5.19	1.50	5.04	1.55
UC3	4.83	1.59	5.08	1.48	4.94	1.55

In all cases, the loadings were over 0.6, ranging from 0.68 to 0.83, providing evidence for unidimensionality. The scale reliability (Cronbach's alpha) was over 0.7 for all dimensions. The composite reliability (CR) was above the threshold of 0.7 and the average variance extracted (AVE) was above the threshold of 0.5, thus giving support for the convergent validity of each dimension.

The correlation between the three dimensions is statistically significant ( $p < 0.001$ ), the correlation coefficient ranging from 0.68 to 0.73.

Table 5. GOF indices

	N	$\chi^2$	DF	$\chi^2/DF$	TLI	CFI	RMSEA
Ro	758	120.54	34	3.55	0.95	0.97	0.058
Ro-M	418	68.43	17	4.025	0.94	0.97	0.085
Ro-F	340	52.11	17	3.065	0.95	0.97	0.078

The model fit with the data is very good. Although the chi-square is significant, the goodness-of-fit (GOF) indices are within the limits recommended by Hair et al. [12]. The testing results are summarized in Table 5

The model testing results on each group provide evidence for the configural invariance of the Romanian sample across gender.

Next step is to test the metric invariance by constraining the loadings to be equivalent. The model comparison shows a nonsignificant chi-square difference ( $\Delta\chi^2 = 2.30$ ,  $\Delta DF = 5$ ,  $p = .806$ ), therefore the model exhibits metric invariance. This means that the model has been perceived in the same way in each group.

Testing the scalar invariance is done by constraining the intercepts to be equivalent. Although the model comparison shows a significant chi-square difference ( $\Delta\chi^2 = 17.55$ ,  $\Delta DF = 8$ ,  $p = 0.004$ ), the depreciation of CFI is less than 0.01 so the model has scalar invariance, according to this criterion [7].

The test for the structural covariance invariance resulted in a non-significant chi-square difference ( $\Delta\chi^2 = 8.94$ ,  $\Delta DF = 6$ ,  $p = 0.685$ ). The results of the invariance analysis are presented in Table 6.

Table 6. Model comparison for the Romanian sample (N=758)

Model	$\Delta\chi^2$	$\Delta DF$	p	CFI	$\Delta CFI$
unconstraint				0.969	
measurement weights	2.30	5	0.806	0.970	0.001
measurement intercepts	17.55	8	0.004	0.964	-0.006
structural covariances	8.94	6	0.685	0.965	0.001

Overall, the invariance analysis results enable an analysis of gender differences for the Romanian sample.

### Lithuanian sample

The descriptive statistics for the sample and each gender is presented in Table 7.

Table 7. Descriptive statistics (N=297)

Item	Male (118)		Female (179)		Total	
	M	SD	M	SD	M	SD
US1	4.30	1.63	4.37	1.66	4.34	1.64
US2	3.78	1.84	3.68	1.92	3.72	1.88
UI1	4.58	1.83	4.49	1.94	4.53	1.89
UI2	4.99	1.60	4.60	1.86	4.75	1.77
UI3	4.25	1.76	3.81	1.87	3.98	1.84
UC1	4.60	1.73	4.66	1.83	4.64	1.78
UC2	4.70	1.62	4.99	1.70	4.88	1.67
UC3	4.79	1.61	4.74	1.67	4.76	1.64

The loadings were over 0.6, with one exception, ranging from 0.56 to 0.90. The correlation between the three dimensions is ranging from 0.68 to 0.90. The scale reliability (Cronbach's alpha) was over 0.7 for all dimensions.

The composite reliability and the average variance extracted were above the threshold of 0.7, respectively 0.5 (except for one dimension), thus providing evidence for the convergent validity of each dimension.

The model fit with the Lithuanian data is acceptable. Although the chi-square is significant, the goodness-of-fit (GOF) indices are within the limits recommended by Hair et al. [12]. The testing results are summarized in Table 8.

The model testing results provide evidence for the configural invariance of the Lithuanian sample across gender.

Table 8. GOF indices

	N	$\chi^2$	DF	$\chi^2/DF$	TLI	CFI	RMSEA
Lt	297	100.30	34	2.950	0.90	0.94	0.081
Lt-M	118	68.42	17	4.025	0.94	0.97	0.085
Lt-F	179	52.11	17	3.065	0.95	0.97	0.078

The test for metric invariance shows a nonsignificant chi-square difference ( $\Delta\chi^2 = 2.30$ ,  $\Delta DF = 5$ ,  $p = .361$ ), therefore the model exhibits metric invariance. Testing the scalar invariance also shows a nonsignificant chi-square difference ( $\Delta\chi^2 = 14.82$ ,  $\Delta DF = 8$ ,  $p = 0.063$ ), so the model has scalar invariance.

The test for the structural covariance invariance resulted in a non-significant chi-square difference ( $\Delta\chi^2 = 3.63$ ,  $\Delta DF = 6$ ,  $p = 0.726$ ). The results of the invariance analysis are presented in Table 9.

Table 9. Model comparison for the Lithuanian sample (N=297)

Model	$\Delta\chi^2$	$\Delta DF$	p	CFI	$\Delta CFI$
unconstraint				0.939	
measurement weights	5.47	5	0.361	0.938	-0.001
measurement intercepts	14.82	8	0.063	0.932	-0.006
structural covariances	3.63	6	0.726	0.924	-0.008

Overall, the invariance analysis results enable an analysis of gender differences for the Lithuanian sample.

### Gender analysis

The gender differences are firstly analyzed at dimension level. The descriptive statistics (mean value and standard deviation) for each group are presented in Table 11.

Table 11. Gender differences in the usefulness dimensions

Gender		Romania			Lithuania		
		US	UI	UC	US	UI	UC
Male	M	4.20	4.63	4.91	4.04	4.61	4.70
	SD	1.40	1.32	1.40	1.56	1.50	1.32
Female	M	4.32	4.80	5.11	4.03	4.30	4.80
	SD	1.36	1.36	1.32	1.59	1.60	1.45

In the Romanian sample, female students have a higher perception of each Facebook usefulness dimension. A one-way ANOVA (1, 756, 757) showed that the gender differences are marginally significant for the information usefulness ( $F=2.895$ ,  $p=.089$ ), and significant for the collaboration usefulness ( $F=3.990$ ,  $p<0.046$ ).

In the Lithuanian sample, male students have a higher perception of the social and information usefulness while female students valued higher the collaboration usefulness. The one-way ANOVA test (1, 295, 296) showed that the differences are marginally significant for the information usefulness ( $F=2.750$ ,  $p<0.098$ ).

A more detailed comparison could be done at observed score level. The mean values are presented in Table 13.

In the Romanian sample, female students scored higher all items. The differences are statistically significant for three items (UI1, UC2, UC3) and marginally significant for two (US1, UI3). In the Lithuanian sample, male students scored higher five items (US2, UC3, and all information usefulness items). However, only one difference was significant (UI3) and another marginally significant (UI2).

Table 12. Means of the observed scores

	US1	US2	UI1	UI2	UI3	UC1	UC2	UC3
Ro-M	4.47	3.93	4.68	4.78	4.43	4.99	4.92	4.83
Ro-F	4.68	3.97	4.96	4.79	4.63	5.07	5.19	5.08
Lt-M	4.30	3.78	4.58	4.99	4.25	4.60	4.70	4.79
Lt-F	4.37	3.68	4.49	4.60	3.81	4.66	4.99	4.74

As regards the structural covariance, the gender differences are small, except for the correlation between the information and collaboration dimensions, which is stronger for the female students in the Lithuanian sample.

### Cross-country analysis

Given the gender differences in each country, a cross-country comparison makes sense for each gender. In order to do this, MGCFA and an invariance analysis have been carried on for each gender.

The model comparison results for the male sample (N=536) shows metric, scalar, and covariance invariance (see Table 13).

Table 13. Model comparison for the male sample (N=536)

Model	$\Delta\chi^2$	$\Delta DF$	p	CFI	$\Delta CFI$
unconstraint				0.962	
measurement weights	14.78	5	0.011	0.956	-0.006
measurement intercepts	15.07	8	0.058	0.953	-0.003
structural covariances	16.94	6	0.009	0.947	-0.006

Romanian male students scored higher on each dimension. The ANOVA test showed that the differences are not statistically significant. At observed score level, Romanian male students scored higher, except for the item UI2 (getting useful information from university people). The ANOVA test showed that the difference is only statistically significant only for the item UC1.

The model comparison results for the female sample (N=519) shows only a metric invariance. The results are presented in Table 14.

Table 14. Model comparison for the female sample (N=519)

Model	$\Delta\chi^2$	$\Delta DF$	p	CFI	$\Delta CFI$
unconstraint				0.959	
measurement weights	9.47	5	0.092	0.957	-0.002
measurement intercepts	30.88	8	0.000	0.945	-0.012
structural covariances	20.91	6	0.002	0.937	-0.008

The cross-country comparison shows that the Romanian female students have a higher perception as regards all items of the Facebook usefulness. The ANOVA test showed that the differences are statistically significant, except for the items US2 and UC2.

## DISCUSSION

The analysis of invariance across gender in two countries provides evidence for the configural, metric, and scalar invariance. This finding has two implications: it brings further evidence for the reliability of the measurement scale and enables gender and cultural analyses.

The main contribution of this study is the analysis of gender differences as regards the Facebook usefulness as perceived by university students from Lithuania and Romania. This analysis has been performed both at country level and across the two countries.

In both countries, both genders perceived Facebook as being more useful for collaboration, then for information.

In Romania, female students have a higher perception of the Facebook usefulness. The gender differences are higher and statistically significant for most of the Facebook usefulness measures. The results are consistent with previous findings showing that females are more interested in the academic use of Facebook [19, 23].

In Lithuania, the results show that male students have a higher perception of information support while female students scored higher than their male colleagues on the collaboration support.

A third finding is that for the female students, Facebook has been perceived as being more useful by Romanians than by Lithuanians. A possible explanation is a difference as regards the size of the Facebook network which is much larger in the Romanian sample. Having more Facebook friends is increasing the information source and provides more opportunities for collaboration. In the case of male students, the cross-country comparison shows a higher perception of the Romanians.

There are several limitations that should be considered for future research. First, there are typical limitations because of the cross-sectional nature of the study. Second, the study is limited by the overall objective of identifying the potential educational outcomes of Facebook. As such, it does not consider the usefulness for socialization and entertainment. Third, the data have been collected from only two countries.

## CONCLUSION AND FUTURE WORK

This study contributes to a better understanding of the perceived usefulness of Facebook for university students. The results provide evidence for the reliability and validity of the scale that exhibits metric and scalar invariance across gender in each country.

There are several research directions to advance the study of gender and cultural differences as regards the Facebook use in universities. Future research needs to validate the model on other samples and check the invariance on samples from other countries. Also, a qualitative study may shed light on each dimension of the Facebook usefulness and explain more specific gender differences. A research question to be answered is if there are gender or cultural differences as regards the information and collaboration usefulness.

## Acknowledgement

This work was supported in part by a national grant financed by ANCS under COGNOTIC 1609 0101 / 2016.

## REFERENCES

1. Anderson, J.C. & Gerbing, D.W. (1988) Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach. *Psych. Bulletin*, 103(3), 411-423.
2. Arbuckle, J. (2006) *AMOS User's Guide*. Amos Development Corporation.
3. Arteaga Sánchez, R., Cortijo, V., & Javed, U. (2014). Students' perceptions of Facebook for academic purposes. *Computers & Education*, 70, 138-149.
4. Balog, A., Pribeanu, C. (2015) A multidimensional model of Facebook usefulness for students. *Revista Romana de Interactiune Om-Calculator* 8(2), 157-172.
5. Brown, J. S., & Adler, R. P. (2008). Open education, the long tail, and learning 2.0. *Educause review*, 43(1), 16-20.
6. Byrne B (2010) *Structural Equation Modeling with AMOS. Basic Concepts, Applications, and Programming*. Lawrence Erlbaum Association.
7. Cheung GW and Rensvold RB (2002) Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modelling* 9(2), 233-255.
8. Donlan, L. (2014). Exploring the views of students on the use of Facebook in university teaching and learning. *Journal of Further and Higher Education*, 1-17.
9. Enoch, Y. and Soker, Z. (2006). Age, gender, ethnicity and the digital divide: University student's use of web-based instruction. *Open Learning: The Journal of Open, Distance, and e-Learning*, 21( 2): 99-110.
10. Eubanks, V. (2000). Paradigms and Perversions: A Women's Place in Cyberspace. *PCSR Newsletter*, 18 (1).
11. Gunn, C. (2003). Dominant or different? Gender issues in computer-supported learning. *Journal of Asynchronous Learning Networks*, 7 (1), 14-30.
12. Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., Tatham, R.L. (2006). *Multivariate Data Analysis*. 6th ed., Prentice Hall, 2006.
13. Junco, R. (2012). The relationship between frequency of Facebook use, participation in Facebook activities, and student engagement. *Computers & Education*, 58(1), 162-171.
14. Kumar, S. (2009) Undergraduate perceptions of the usefulness of Web 2.0 in higher education: Survey development. *Proc. European Conference on e-Learning*, 308-314.
15. Lamanauskas V, Slekiene V, Balog A, Pribeanu C (2013) Exploring the usefulness of social networking websites: a multidimensional model. *Studies in Informatics and Control*, 22(2), 175-184.
16. Lampe C., Vitak, J., Gray, R., Ellison N.B. (2012) Perceptions of Facebook's value as information source. *Proceedings of CHI 2012*, 3195-3204, ACM.
17. Link, T.M., Marz, R. (2006). Computer literacy and attitudes towards e-learning among first-year medical students. *BMC Medical Education*, 6: 34-42.
18. Losh, S. C. (2003). Gender and educational digital chasms in computer and internet access and use over time: 1983-2000. *IT & Society*, 1(4), 73-86.
19. Mazman, S. G., Usluel, Y. K (2011) Gender differences in using social networks TOJET: *The Turkish Online Journal of Educational Technology* 10(2), 133-139.
20. McAndrew, F. T., & Jeong, H. S. (2012). Who does what on Facebook? Age, sex, and relationship status as predictors of Facebook use. *Computers in Human Behavior*, 28(6), 2359-2365.
21. Park, N., Kee, K. F., Valenzuela, S. (2009). Being immersed in social networking environment: Facebook groups, uses and gratifications, and social outcomes. *CyberPsychology & Behavior*, 12(6), 729-733.
22. Rodriguez-Hoyos, C., Haya Salmon, I., Fernandez-Díaz, E. (2015). Research on SNS and education: the state of the art and its challenges. *Australasian Journal of Educational Technology*, 31(1), 100-111.
23. Selwyn, Neil (2008). An investigation of differences in undergraduates' academic use of the internet. *Active Learning in Higher Education* 9(1), 11-22.
24. Steenkamp JBE and Baumgartner H. (1998) Assessing measurement invariance in cross-national consumer research. *Journal of consumer research* 25(1):78-107.
25. Suriya, M., Panteli, A. (2000). The Globalization of Gender in IT. *Women, Work and Computerization: Charting a Course to the Future*. Vancouver BC: Kluwer Academic Publishers.
26. Sutton, R. (1991). Equity and Computers in the Schools: A Decade of Research. *Review of Educational Research* 61(4), 475-503.
27. Vandenberg, R.J., Lance, C.E. (2000). A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. *Organizational Research Methods* 3(1), 4-70.
28. Yates, S.J. (2001). Gender, Language, and CMC for Education. *Learning and Instruction* 11, 21-34.

# The social media impact on local e-government institutions in Romania

Maria Gheorghe-Moisii, Tirziu Eugenia, Dragos Daniel Iordache,  
ICI Bucharest

Bd. Mareşal Averescu nr.8-10, Bucharest  
[moise@ici.ro](mailto:moise@ici.ro), [ginet@ici.ro](mailto:ginet@ici.ro), [iordache@ici.ro](mailto:iordache@ici.ro)

## ABSTRACT

In this research paper we analyze the extent and manner in which e-government institutions from public administration currently use social media networks. The sample consist in a number of 30 local e-government institutions analyzed using the collective case study, based on a set of indicators for three social media networks, frequently used in Romania. The results showed a larger representation of Romanian local e-government institutions on Facebook and less on Twitter or YouTube.

## Author Keywords

Social media, e-government, local e-government institutions, case study.

## ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

## General Terms

Human Factors

## INTRODUCTION

Social media networks gained a remarkable popularity in recent years. They are currently key communication tools. E-government institutions can benefit from social media facilities to improve their communications. Social media networks can help promote public institutions' initiatives and inform citizens and beneficiaries of social services. At the same time, social media networks have the advantage of capitalizing on the dialogue with the public and improving the internal communication processes of the institutions.

Social Media networks allows real-world communication in both ways, meaning not only the use of special tools or techniques, but also the promotion of digital engagement through the transition from disseminating information (site) to conversation, from an observer to listening and to respond, from public awareness to pro-activity. At the same time, the social networking level can create an online learning center. Online Learning Centers can be used in e-government as an online training and education infrastructure. Boxer and Johnson (2002) suggest that the Learning and Training Center should serve as a focal point for all government learning activities.

The digital commitment that an institution or public authority assumes facilitates community involvement in decision-making at the local level.

This study identifies the key impact analysis of social media networks as instruments to ensure greater transparency of public institutions, increase awareness and

participation of citizens in decision-making and improved quality of services.

The following section presents related work in the area of the impact of social media networks in e-government. In the next section the evaluation results are presented and discussed. The paper ends with conclusion and directions for future work.

## RELATED WORK

The impact of social media networks in e-government has been studied from different theoretical perspectives by using various methods. Few approaches exist that investigate the impact of social media networks in e-government using case study perspective in Romania.

Bonsón et al. (2015) conducted a study in order to measure the impact of social media and content types on stakeholders' engagement on Western European local governments' Facebook pages. The sample includes 15-member countries of the EU via 75 local governments belonging to four different public administration styles, analyzing 50 posts from each municipality. The posts were coded into 16 content types (e.g. environment, housing) and 5 media types (e.g. video, text), respectively. Metrics of popularity, commitment and engagement were also computed for each post. Results show that the most used media types are links and photos. Cultural activities, sports and marketing related topics are the most widely posted by Western European local governments, but these contents do not seem to be the most relevant for citizens. Local governments' preferences to different media and content types seem to depend on the institutional context. Significant differences on engagement levels by citizens have been found among media and content types and in different institutional contexts.

The study conducted by Mickoleit et al (2014) takes a comparative analysis of social media use in and by Organization for European Economic Cooperation (OECD) governments. The analysis is based on a large amount of empirical data, including a survey of OECD governments on policies and objectives in this area. The paper proposes tools to guide decision makers: a checklist of issues to be considered by government institutions, a set of potential indicators to appraise impacts, and a range of options for more in-depth policy analysis. The set of potential indicators to appraise impacts were mainly related to presence, popularity, penetration, perception and purpose.

The study of Lev-On & Steinfeld (2015) analyzed the scope and character of the activities of the Israeli



municipalities on Facebook. The study maps the Facebook presence of Israeli municipalities and checks for correlations between Facebook presence, municipality size and the geographic and socio-demographic characteristics of the municipalities. Several engagement indices of municipal activities on Facebook (number of fans, average likes, comments, and shares of posts) have been measured and checked for correlations of these engagement indices with municipality size and geographic and social demographic characteristics of the municipalities. The study found that municipalities' Facebook adoption is significant. More than half of the Israeli municipalities maintained an official Facebook page at the time of data collection.

Hrdinova et al. (2010) developed a framework of 8 elements that can be addressed for a successful social media policy: employee access, account management, acceptable use, employee conduct, content, security, legal issues, and citizen conduct.

## CASE STUDIES

### Method and procedure

As a method of collecting and analyzing data, we used the collective case study, structured on the basis of a set of indicators for each social network analyzed for assessing social media impact in e-government. The case study can be defined as the research that provides a detailed analysis and explanation of one or more cases (Johnson & Christensen, 2008). According to Stake (1995), three types of case studies can be identified: intrinsic, instrumental and collective. In the collective case study, the researcher believes that he can reach a much deeper understanding of the subject of the research by concomitantly studying several cases in a comprehensive research.

Three social media networks: Facebook, Twitter, and YouTube have been analyzed. The presence of each municipality was examined by the official name provided on their site. In Facebook, we checked the following: the existence of an official Facebook page, the number of visitors, the number of followers, the number of likes, the number of conversations, the number of ratings and the engagement score.

In Twitter, we've checked the following: existence of an official Twitter account, number of Twitter tweets, number of Twitter following, number of Twitter followers, number of twitter rating and number of Twitter lists. In YouTube, we analyzed the following: existence of an official YouTube channel, number of subscribers to the official YouTube channel, YouTube Chat Counts, and YouTube Interactions as Response to Comments.

To analyze the data obtained through site content analysis, an exploratory analysis was first conducted to provide a broad perspective on how to use social media networks. Moreover, each local government has been assessed according to its active presence.

In the present study, social media impact assessment in e-government has three main objectives: inventory, assessment and guidance to decision makers. To achieve these objectives, both quantitative and qualitative data are required, but at different degrees in each phase. It has been

demonstrated that Facebook's social media has the biggest impact on citizens in Romania, being the most visited and most official sites, as shown in Figure 1.

The development of case studies involved, in the first instance, a hierarchical classification of the objectives that provide a mature preliminary criterion for measuring social media impact in e-government (Table 1).

Social Media	Indicators
FACEBOOK	<ul style="list-style-type: none"> <li>▪ number of visitors</li> <li>▪ number of followers</li> <li>▪ number of likes</li> <li>▪ number of ratings</li> <li>▪ engagement score</li> <li>▪ number of conversations.</li> </ul>
TWITTER	<ul style="list-style-type: none"> <li>▪ number of Twitter tweets</li> <li>▪ number of Twitter following</li> <li>▪ number of Twitter followers</li> <li>▪ Twitter rating</li> <li>▪ number of Twitter lists</li> </ul>
YOUTUBE	<ul style="list-style-type: none"> <li>▪ number of subscribers</li> <li>▪ interactions as response to comments</li> <li>▪ uploads on YouTube during the evaluation period</li> </ul>

**Table 1. Indicators used in Social media assessment in e-government**

These include the indicators analyzed to test their feasibility and relevance for comparisons, for example on the presence, perception and purpose of using social media in governance.

### Sample

In order to assess the impact of social media networks on e-government institutions in Romania, empirical research was carried out based on a sample of 30 public administration sites, taking into account the comparison of different social media networks.

The sample used was selected according to the number of inhabitants of the cities. The analysis of the use of government websites by means of social media was conducted between April and May 2017. First of all, the presence of town halls in the virtual space of social media was analyzed. The second part of the research consisted of analyzing social media platforms. In addition to the elements included in the official site, the presence and activity of local governments from the most widely known social media platforms were measured using different indicators such as number of followers, number of comments, number of likes, and number of shares.

## RESULTS

### Evaluation of the impact on Facebook network

Facebook page analysis was done using LikeAlyzer instrument. LikeAlyzer measures and analyzes the potential and effectiveness of the Facebook page. It monitors and compares all the features of the Facebook page.

In the Figure 1 we can see the presence and the involvement of local e-government institutions in Romania that have been analyzed in this study. The generated results have highlighted a major gap between the sites analyzed. Facebook is the most preferred network in terms of as presence, interaction and information distributed by local e-government institutions in Romania.

From the 30 local e-government institutions, a number of 27 institutions have an official page on Facebook. Three

institutions do not have an official Facebook page, but there are unofficial pages create by users.

The number of the followers varied between 20412 and 54 with a mean of 5565. There are many who visit sites, just to inform themselves, but they do not actually interact. Therefore, the engagement score is relevant, measuring the interaction of citizens.

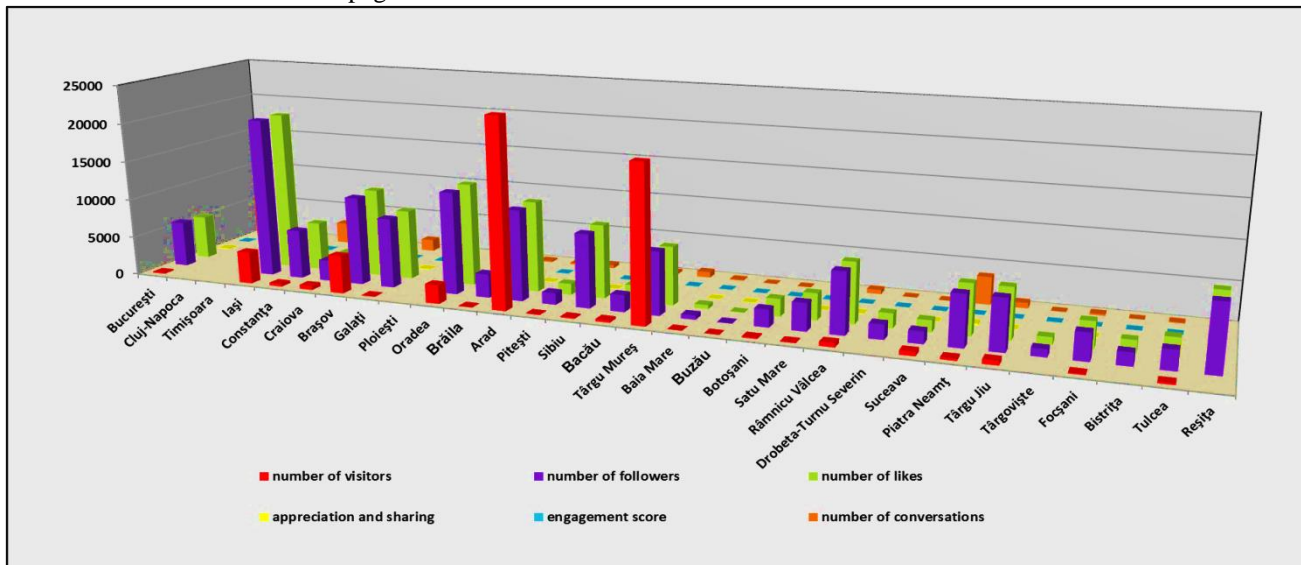


Figure 1: The presence and involvement of local e-government institutions on Facebook

### Evaluation of the impact on Twitter network

The presence and the involvement of local e-government institutions analyzed in this study are presented in Figure 2. The analysis of Twitter pages was done by analyzing official public websites. The results highlighted a minimal involvement of the local e-government institutions on the Twitter network. Only seven of the 30 municipalities

surveyed have an official Twitter account but also a weak official data.

The number of the tweets varied between 6089 and 8 with a mean of 2627. Five of the seven online present have tweets and all of the sites have followers. Therefore the Twitter network is very little used in e-government by local e-government institutions in Romania.

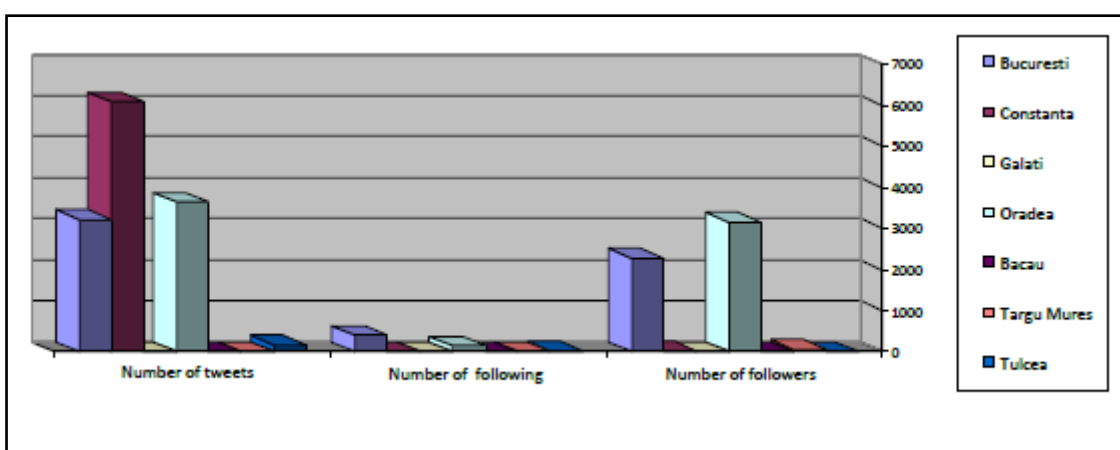


Figure 2: The presence and involvement of local e-government institutions on Twitter

### Evaluation of the impact on YouTube network

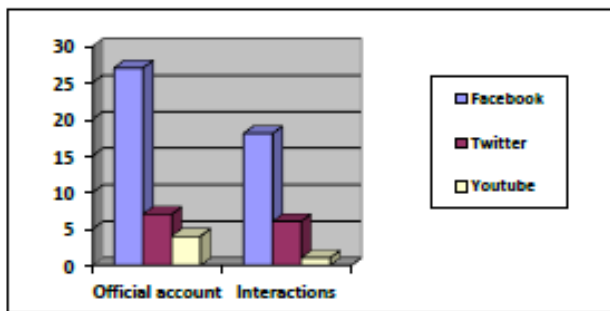
The evaluation was carried out by analyzing the official public sites. The generated results highlighted a major discrepancy in the presence and media coverage of online

information of institutions, resulting in a minimal interaction impact of this network on the sample used. Only four of the 30 municipalities have an official channel, but also that the information provided rather

summarily. The number of uploads on YouTube during the evaluation period varied between 18 and 2 with a mean of 9 uploads.

### Comparative results on the three Social Media networks

A comparative study of the three networks underlined highlights the fact that the Facebook network has the highest share in presence, transparency and collaboration. The local e-government institutions in Romania are mostly represented in the online environment by the Facebook network. Thus, the institution's reporting to the Facebook network is the most relevant (see Figure 3).



**Figure 3: Social Media Comparison of Online Interaction**

### CONCLUSION AND FUTURE WORK

The impact of social media on e-government processes can be derived from the openness and transparency of e-government to the public, the collaboration between different administrative units in the process of public policy making, participation of citizens in political decisions, faster dissemination and greater information.

It is important for local e-government institutions to have a presence on social media sites in order to build a high level of people trust especially if governments use social media to provide citizens with an active information channel to ensure that governments not only listen or monitor the process but also respond.

Social media networks give local e-government institutions an opportunity to take the advantage of using more creative and innovative ways in communicating with

citizens and making effective use of the valuable resources available on such sites and outside its control. Future research may have the following main working directions:

- Carrying out intercultural studies for the comparative analysis of the integration of social media networks in different e-government applications;
- Developing software applications for monitoring social media impact in e-government.

### ACKNOWLEDGMENTS

This work has been supported by two Romanian grants, financed by ANCS under COGNOTIC 1609 0101 / 2016 and COGNOTIC 1609 0601 / 2017.

### REFERENCES

1. Bonsón, E., Royo, S., & Ratkai, M. Citizens' engagement on local governments' Facebook sites. An empirical analysis: The impact of different media and content types in Western Europe. *Government Information Quarterly*, 32(1), (2015), 52-62.
2. Boxer, K. M., & Johnson, B. How To Build an Online Learning Center. *T+D*, 56(8), (2002), 36-42.
3. Hrdinova, J.; Helbig, N.; Peters, C.S. Designing social media policy for government: Eight essential elements. *Center for Technology in Government*: Albany, NY, USA, 2010.
4. Johnson, B., & Christensen, L. Educational research: Quantitative, qualitative, and mixed approaches. *Sage*, 2008
5. Lev-On, A., & Steinfeld, N. Local engagement online: Municipal Facebook pages as hubs of interaction. *Government Information Quarterly*, 32(3), (2015), 299-307.
6. Mickoleit, A. "Social Media Use by Governments: A Policy Primer to Discuss Trends, Identify Policy Opportunities and Guide Decision Makers", *OECD Working Papers on Public Governance*, No. 26, (2014), OECD Publishing <http://dx.doi.org/10.1787/5jxrcmgmhmk0s-en>
7. Stake, R. E. The art of case study research. *Thousand Oaks, CA: Sage*, 1995
8. <http://likealyzer.com>.

# Identifying Similarities between Tweets

**Iuliana Minea**

Faculty of Computer Science,  
“Alexandru Ion Cuza”

University of Iasi

General Berthelot, No. 16  
iuliana.minea@info.uaic.ro

**Adrian Iftene**

Faculty of Computer Science,  
“Alexandru Ion Cuza”

University of Iasi

General Berthelot, No. 16  
adiftene@info.uaic.ro

## ABSTRACT

In this paper, an application built by us with aim to provide users the possibility to explore topics from Twitter and find out people’s opinion even if it is a positive one or a negative one is presented. The result of searched topic was divided into groups, based on the named entity from tweets. Using algorithms that calculate distance between two strings, similar tweets was removed. More than that, users have the opportunity to see a sentiment analysis of tweets.

## Author Keywords

Twitter; information retrieval; strings similarity.

## ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI):  
Miscellaneous.

## General Terms

Human Factors; Design.

## INTRODUCTION

In 2012, more than 100 million users posted 340 million tweets a day and the service handled an average of 1.6 billion search queries per day [20]. In 2013, it was one of the ten most-visited websites and has been described as “the SMS of the Internet” [2]. As of February 2017, Twitter had more than 313 million monthly active users [19]. On the day of the 2016 U.S. presidential election, Twitter proved to be the largest source of breaking news, with 40 million tweets sent by 10 p.m. that day [13].

Every day, millions of people use Twitter to create, discover and share ideas with other. Now, people are turning to Twitter as an effective way to reach out to businesses, too. From local stores to big brands, and from brick-and-mortar to internet-based or service sector, people are finding great value in the connections they make with businesses on Twitter [1].

There are many great business uses for Twitter, like sending out news briefs or advertising the latest job opening. But believe it or not, there are even more personal uses for Twitter. With this round up, consider the seven that follow [9].

## SIMILAR APPLICATIONS

**IceRocket** [4] is generally for blog searches, but it offers the possibility to search news on Twitter. Twitter search

returns most recent tweets that relate to your query. If the query is also a user, it shows a fact box about the user, along with tweets by that user. This application lets you reply directly to the tweets. The IceRocket site is a free resource for people looking to monitor their brand, it is ad supported. IceRocket has an API that it licenses to social media monitoring firms as well as PR agencies [21].

**Twitonomy** [18] is an online platform and in order to use it you have to connect with your Twitter account. The user has the opportunity to monitor his account or any other Twitter user, along with lists and any keyword search he wants to watch. Twitonomy acts like a basic Twitter client. You can reply, retweet and favorite individual tweets. As can be observed in image 2 from below, this application has all kinds of statistics about your own activity. You can see things like: tweets per day, how many links you’ve shared, how often others mentioned you in their tweets, which of your tweets were retweeted the most [10].

## APPLICATION DESCRIPTION

This application aims to offer users the possibility to follow news on Twitter, without having to read duplicate topics (this paper represent an extended version of results presented in [5, 11]). In order to remove similar news, it was used a similarity algorithm which calculates distance between two tweets. In the first phase, it was made an analysis, in terms of time and accuracy, of the next four similarity distance algorithms: *Levenshtein*, *Needleman-Wunsch*, *Jaro-Winkler* and *Smith-Waterman*. Upon review of the analysis result, Smith-Waterman turned out to be more competent to find similarities.

When the user performs a search in application, it is done a request to Twitter API in order to retrieve the latest and the most popular tweets. The search result was divided into subcategories. Each category was extracted from tweets and can be: *a location*, *a person*, *an organization* or *a date*. Also tweets was analyzed from the sentiment point of view. A tweet can be categorized as *positive*, *negative* or *neutral*.

## Similar Algorithms

In computer science and statistics, the **Jaro–Winkler distance** is a measure of similarity between two strings [6, 22]. It is a variant of the **Jaro distance** metric a type of string edit distance, and was developed in the area of record linkage (duplicate detection). The lower the Jaro–Winkler

distance for two strings is, the more similar the strings are. The similarity score is normalized such that 0 equates to no similarity and 1 is an exact match. The Jaro score  $d_j$  of two given strings  $s_1$  and  $s_2$  is:

$$d_j = \begin{cases} 0, & \text{if } m = 0 \\ \frac{1}{3} \left( \frac{m}{|s_1|} + \frac{m}{|s_2|} + \frac{m-t}{m} \right), & \text{otherwise} \end{cases}$$

where  $m$  is the number of matching characters,  $t$  is half the number of transpositions (the number of matching characters (but different sequence order) divided by 2).

Jaro-Winkler distance uses a prefix scale  $p$  which gives more favorable ratings to strings that match from the beginning for a set prefix length  $l$ . Given two strings  $s_1$  and  $s_2$ , their Jaro-Winkler distance  $d_w$  is:

$$d_w = d_j + (lp(1 - d_j))$$

where:  $d_j$  is the Jaro distance for string  $s_1$  and  $s_2$ ,  $l$  is the length of common prefix at the start of the string up to a maximum of 4 characters,  $p$  is a constant scaling factor for how much the score is adjusted upwards for having common prefixes (the standard value for this constant is Winkler's work is  $p = 0.1$ ).

The **Levenshtein distance** is a string metric for measuring the difference between two sequences [8]. Informally, the Levenshtein distance between two words is the minimum number of single-character edits (i.e. insertions, deletions or substitutions) required to change one word into the other. Mathematically, the Levenshtein distance between two strings  $a$ ,  $b$  (of length  $|a|$  and  $|b|$  respectively) is given by  $lev_{a,b}(|a|, |b|)$  where:

$$lev_{a,b}(i, j) = \begin{cases} \max(i, j) & \text{if } \min(i, j) = 0 \\ \min \begin{cases} lev_{a,b}(i-1, j) + 1 \\ lev_{a,b}(i, j-1) + 1 \\ lev_{a,b}(i-1, j-1) + 1_{a_i \neq b_j} \end{cases} & \text{otherwise} \end{cases}$$

where  $1_{a_i \neq b_j}$  is the indicator function equal to 0 when  $a_i = b_j$  and equal to 1 otherwise,  $lev_{a,b}(i, j)$  is the distance between the first  $i$  characters of  $a$  and the first  $j$  characters of  $b$ .

The **Needleman-Wunsch algorithm** is an algorithm used in bioinformatics to align protein or nucleotide sequences [7, 12]. It was one of the first applications of dynamic programming to compare sequences. The Needleman-Wunsch algorithm, which is based on dynamic

programming guarantees finding the optimal alignment of pairs of sequences. The algorithm essentially divides a large problem (e.g. the full sequence) into a series of smaller problems and uses the solutions to the smaller problems to reconstruct a solution to the larger problem.

In order to perform a Needleman-Wunsch alignment, a matrix is created which allows us to compare the two sequences. The first line and the first column from the matrix was initialized with line index, respectively column index. The score of the rest cells was calculated as follows:

$$M[i][j] = \min \begin{cases} M[i-1][j] + gap \\ M[i][j-1] + gap \\ M[i-1][j-1] + cost \end{cases}$$

where:  $gap$  is the gap penalty (in this implementation the gap has value 2),  $cost$  was 0 if the characters matches and 1 otherwise. The Needleman-Wunsch distance has a value between 0 (identical strings) and 1 (different strings), and is calculated based on the next formula:

$$D = \frac{M[|s_1|][|s_2|]}{\max(|s_1|, |s_2|) * gap}$$

where  $|s_1|, |s_2|$  is the length of  $s_1$ , respectively  $s_2$  string.

The **Smith-Waterman algorithm** is a dynamic programming method for determining similarity between nucleotide or protein sequences [16]. The algorithm was first proposed in 1981 by Smith and Waterman and is identifying homologous regions between sequences by searching for optimal local alignments. The Smith-Waterman algorithm is built on the idea of comparing segments of all possible lengths between two sequences to identify the best local alignment. It is based on calculation of local alignments instead of global alignments of the sequences and allowing a consideration of deletions and insertions of arbitrary length [16]. The Smith-Waterman algorithm is the most accurate algorithm when it comes to search databases for sequence homology, but it is also the most time consuming [16].

### Named Entity Recognizer Module

This module deals with extracting information, localize and classify named entities in tweets into pre-defined categories such as the names of persons, organizations, locations, expressions of times, etc. For all these operations it was used "Stanford Named Entity Recognizer (NER)" library with 7 class model.

Named Entity Recognition (NER) labels sequences of words in a text which are the names of things, such as person and company names, or gene and protein names. It comes with well-engineered feature extractors for Named



Entity Recognition, and many options for defining feature extractors. Stanford NER is also known as CRFClassifier. The software provides a general implementation of (arbitrary order) linear chain Conditional Random Field (CRF) sequence models [17].

### Tweets Classifier Module

Formerly known as Twitter Sentiment, Sentiment140 [15] is a service that lets users discover the current sentiment around a brand, product or topic on Twitter. Sentiment140 uses classifiers based on machine learning algorithms and allow users to see the classification of individual tweets. The API lets users classify tweets and integrate sentiment analysis classifier functionality into their own websites or applications. The API uses RESTful calls and responses are formatted in JSON. As can be seen in the table 3, the response is the same as the request, except a new field “polarity” added to each object. The polarity values are: 0: *negative*; 2: *neutral*; 4: *positive* [14].

The approach is to use different machine learning classifiers and feature extractors. The machine learning classifiers are Naive Bayes, Maximum Entropy (MaxEnt), and Support Vector Machines (SVM). The feature extractors are unigrams, bigrams, unigrams and bigrams, and unigrams with part of speech tags. We build a framework that treats classifiers and feature extractors as two distinct components [3].

The emoticons are stripped out from training data. If the emoticons are left in, there is a negative impact on the accuracies of the MaxEnt and SVM classifiers, but little effect on Naive Bayes. The difference lies in the mathematical models and feature weight selection of MaxEnt and SVM. Stripping out the emoticons causes the classifier to learn from the other features (e.g. unigrams and bigrams) present in the tweet. The classifier uses these non-emoticon features to determine the sentiment [3].

### Best Similarity Algorithm

In the first phase of application development was aimed to select the best similarity algorithm. The four algorithms introduced above were applied on a set of 2,000 tweets, tweets which were saved in XML files. For detecting which tweets from the entire source data are similar, each tweet was compared with all that follow it.

In the next table can be observer how long did the execution of each algorithm. In order to improve the execution time, it was used a caching mechanism from Microsoft. In cache was kept the comparison result for every algorithm and the key format is: “AlgorithmName\_\*TextTweetOne\_\*TextTweetTwo”.

Another way to improve the comparisons accuracy and time was to remove the stop words (most common words like “the”, “and”, “an” etc.) from tweets. After implementing this step and with the help of caching mechanism, the time was improved (see Table 1).

	Without cache	With cache	With cache after removing stop words
Jaro-Winkler	06:05.35	03:50.11	01:46.06
Levenshtein	22:35.37	12:10.08	11:26:82
Needleman-Wunsch	51:11.47	34:21.63	21:22.49
Smith-Waterman	39:46.32	35:33.20	23:51.46

Table 1. Algorithms execution time.

Three of the algorithms find 505 tweets with distance 0 (tweets are similar), while Smith-Waterman find 569 tweets with distance 0 (it sees a tweet and a retweet being the same). In the end, we decided to use Jaro-Winkler algorithm to group similar tweets due to its fast running time.

In the application interface can be observed how the groups are display and the tweets that were classified as positive have a green background and the negative ones have a red background (see Figure 1).

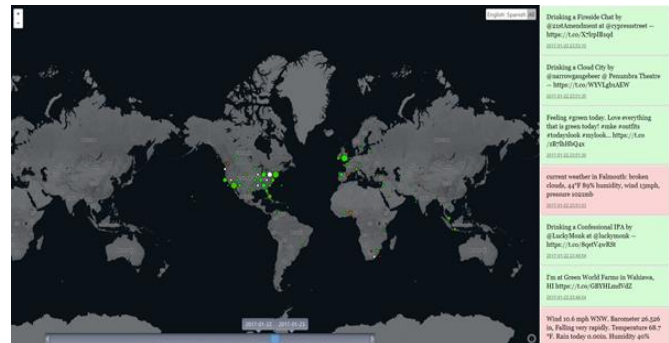


Figure 1. Viewing clusters on Google Maps.

### CONCLUSION

Although there are some applications which extract tweets and perform a sentiment analysis (“Social Mention”, “Twitter Sentiment Visualization”), “News on Twitter” application keeps those benefits and provides other functionality: tweets which contain your searched data, but also grouped by relevant information like: *date*, *location*, *person* and *organization*. Having these clusters, user can decide to read only some topics and not all the tweets returned by API.

When a tweet contains a link, the url is altered to 23 characters even if the link itself is less than 23 characters long. So it is possible to have two tweets that refer to the same thing, have the same destination url, but the links are displayed differently. Because of this thing, Smith-Waterman algorithm can find these two tweets as different.

In order to improve the accuracy of the algorithm, before a comparison, the links can be removed from tweets.

Due to the modular structure, the application functionality can be very easily extended and provide functionality like: which tweets are most retweeted, monitor tweets from users etc.

#### ACKNOWLEDGMENT

This work is partially supported by POC-A1-A1.2.3-G-2015 program, as part of the PrivateSky project (P\_40\_371/13/01.09.2016).

#### REFERENCES

1. Ask a Ronlee: <http://askaaronlee.com/10-reasons-why-your-business-should-use-twitter/> (Last accessed on 29 May 2017)
2. Business Standard: [http://www.business-standard.com/article/technology/swine-flu-s-tweet-tweet-causes-online-flutter-109042900097\\_1.html](http://www.business-standard.com/article/technology/swine-flu-s-tweet-tweet-causes-online-flutter-109042900097_1.html) (Last accessed on 29 May 2017)
3. Go, R., Bhayani, L. and Huang, L. Twitter Sentiment Classification Using Distant Supervision. *Stanford University*, Technical Paper (2009).
4. IceRocket: <http://www.icerocket.com/> (Last accessed on 29 May 2017)
5. Iftene, A., Dudu, M. Ş. and Miron, A. R. Scalable system for opinion mining on Twitter data. Dynamic visualization for data related to refugees' crisis and to terrorist attacks. In *Proceedings of 26<sup>th</sup> International Conference on Information System Development, ISD 2017*, September 6-8, Larnaca, Cyprus (2017).
6. Jaro. M. A. Advances in record linkage methodology as applied to the 1985 census of Tampa Florida. *Journal of the American Statistical Association*, 84 (406), (1989), 414-420.
7. Lesk, A. *Introduction to Bioinformatics*. Oxford University Press (2002).
8. Levenshtein, V. I. Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*. 10 (8), (1966), 707-710.
9. Lifewire: <https://www.lifewire.com/ways-for-beginners-to-use-twitter-3486595> (Last accessed on 29 May 2017)
10. Marketingland Twitonomy: <http://marketingland.com/twitonomy-the-best-twitter-analytics-tool-youve-never-heard-of-55389> (Last accessed on 29 May 2017)
11. Minea, I. and Iftene, A. News on Twitter. In *Proceedings CSM4 (The Fourth Conference of Mathematical Society of the Republic of Moldova)*, June 25 - July 2, Chisinau, Republic of Moldova (2017), 535-538.
12. Needleman, S. B. and Wunsch, C. D. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology*, 48 (3), (1970), 443-453.
13. New York Times: <https://www.nytimes.com/2016/11/09/technology/for-election-day-chatter-twitter-ruled-social-media.html> (Last accessed on 29 May 2017)
14. Programmable Web: <https://www.programmableweb.com/api/sentiment140> (Last accessed on 29 May 2017)
15. Sentiment 140: <http://help.sentiment140.com/api> (Last accessed on 29 May 2017)
16. Smith, T. F. and Waterman, M. S. Identification of Common Molecular Subsequences. *Journal of Molecular Biology*, 147 (1981), 195-197.
17. Stanford CRF: <http://nlp.stanford.edu/software/CRF-NER.shtml> (Last accessed on 29 May 2017)
18. Twitonomy: [http://www.twitonomy.com/?gclid=CO-gi\\_fZ8dECFQuMGQodbgYCKQ](http://www.twitonomy.com/?gclid=CO-gi_fZ8dECFQuMGQodbgYCKQ) (Last accessed on 29 May 2017)
19. Twitter Company: <https://about.twitter.com/company> (Last accessed on 29 May 2017)
20. Twitter: <https://blog.twitter.com/2012/twitter-turns-six> (Last accessed on 29 May 2017)
21. Wikipedia IceRocket: <https://en.wikipedia.org/wiki/IceRocket> (Last accessed on 29 May 2017)
22. Winkler, W. E. String Comparator Metrics and Enhanced Decision Rules in the Fellegi-Sunter Model of Record Linkage. *Proceedings of the Section on Survey Research Methods*. American Statistical Association (1990), 354-359.

# Identification of Agreement and Disagreement in Chat Conversations

**Rasid Cetin**

Politehnica University of Bucharest  
313 Splaiul Independentei,  
Bucharest, Romania  
cetinrasid3@gmail.com

**Ștefan Trăușan-Matu**

University Politehnica of Bucharest  
313 Splaiul Independentei,  
Bucharest, Romania  
and  
Research Institute for Artificial Intelligence  
and  
Academy of Romanian Scientists  
stefan.trausan@cs.pub.ro

## ABSTRACT

A recent pedagogical approach is computer-supported collaborative learning (CSCL) where by means of collective interaction through collaborative interfaces on internet, learning is accomplished by joint knowledge building. One of the most relevant tools that a CSCL uses is chat, which is an effective way to transfer information and collaborate. The purpose of this paper is to offer a technique of analyzing the composition of chats with a design to find the points of convergence and divergence between participants. The points of convergence that are detected in a chat discussion could emphasize a moment where a fact is validated, or a better solution for a problem is discovered or a new concept was assimilated by someone. Going in the same direction, the points of divergence also provide valuable information, because they capture moments where a fact is invalidated or a new idea about an existing topic is introduced. The starting point of this method has its origins in Bakhtin's dialogism and Trausan-Matu's polyphonic model.

**KEYWORDS:** CSCL; chat; convergence; divergence; polyphony; dialogism; pattern

## ACM Classification Keywords

I.2.7 Natural Language Processing: Text analysis.

## INTRODUCTION

Dialogue is an exchange of ideas and opinions on a particular issue, and its structure resembles that of a game, with turns and utterances (the unit of analysis), where every participant waits for his turn to say something, or to play [1]. An utterance is the act of saying. For example, if we have a sentence "Today it is not tomorrow", and someone read it with loud voice two times, he emits two utterances of the same sentence. An utterance is an emitted "piece of language", which has a moment of silence from the utterer before and after it. The essence of an utterance is given by the fact that it has a time, a place, a speaker, a language in which the message transmitted is encoded, but has no special form or content.

Computer-supported collaborative learning (CSCL) is a recent pedagogical approach on which, by means of collective interaction through collaborative interfaces on internet, for example chat (instant messaging), learning is accomplished by joint knowledge building [2]. One advantage of instant messaging is that the participants may introduce utterances at very close moments, even concomitantly, thus a higher degree of interaction is achieved than in face-to-face conversation [3].

System validation is done by using a corpus of chat conversations performed by students. The corpus consists in chat conversations between more than three students. They had to find advantages and disadvantages for how chats, blogs, forums, wiki, and google wave help in collaborative learning. The conversation had to be divided in two parts. The first part consisted in a debate, where to every student is assigned one technology and they had to persuade the others that their technology is the best choice and in the same time to emphasize the disadvantages of using the other technologies. In the second part, the students should reach to a common point [4]. Because of this format of the discussion, various points of convergence and divergence should arise between participants.

The idea of collaborative learning is that a group of people discuss for trying to jointly understand new topics or to find meanings to some concepts or to find solutions to some problems. One advantage is the appearance of controversies, which marks the fact that some participants have conflicting ideas with others. These moments are beneficial because some flaws in the way of thinking are detected and they may be corrected. Another advantage is that because of the high degree of interaction, new ideas are generated and shared between the participants, thus the construction of knowledge is enabled.

There are several applications of the approach presented in this paper. First, it may enhance the collaboration in the classroom by helping the teacher to manage the students. By using this program, the students can have their individualized learning plan that is created by the teacher after keeping track

of their achievements, of how far they are from accomplishing the targets that were set, taking account of their involvement inferred from their agreements and disagreements. Also, collaboration between students can be improved through this technology, because the teacher can group the students using the application, avoiding, for example, to put too many students in the same group if they have a profile with frequent disagreements. An idea that should be investigated in the future is that groups should have students with balanced agreement/disagreement profiles.

In the next section is presented the theoretical background which consists in Trausan-Matu's polyphonic model [3, 8, 9], based on Bakhtin's ideas about dialogism and polyphony [7]. It is also presented the concept of points of divergence and convergence. In the third section are presented some implementation details of the system. In the fourth section are shown some results of the actual system and in the last section are drawn some conclusions.

## STATE OF THE ART

### The Polyphonic Model

Dialogue is a result of a mental and verbal effort which appears when a group of people tries to endorse their way of thinking by providing personal assertion offered as evidence. In order to give voice to their ideas, each participant attempts to choose the right words. Commonly, the dialogue resulted from an educational interaction is constrained by certain restrictions that guides the discussion and not let it to divert from the main topics.

In Bakhtin's theory, the dialogue (as any language communication, including texts) is multivocal [6, 7], it appears as a strong interaction between different perspectives, which are pictured into words and which reflects the participants. Usually, some words are repeated during a dialogue in order to emphasize a particular idea. As a result, starting from Bakhtin's ideas and following the polyphonic method of analysis [6] the main topics or concepts considered in the dialogue can be tracked by finding the most repeated words [6].

Through a dialogical point of view, everything in a chat is connected. Every participant has free will in what he/she says and an individual path in the discussion. However, as Bakhtin wrote, every utterance has echoes of the preceding ones, so they are related all together [7]. An utterance is in stronger relations with recent past utterances to which responds and with future ones whose response it anticipates. Threads of ideas are generated that exist in parallel and that inter-animate according to divergent and convergent patterns [9]. In other words, in chats appear phenomena that resemble with the polyphonic model, in which several voices with individual personalities inter-animate according to dissonant and consonant chords, the first ones driving to the latter.

### Convergence and Divergence

In the CSCL dialogues, students will try to persuade the other students that his ideas and opinions are the most suitable for the given situation. Each one of them will materialize their way of thinking into words and focus on the topic received. Therefore, this interaction will create a transfer of ideas between students, and points of convergence and divergence will appear when participants try to prove their views [9].

The points of convergence depict moments in the conversation where a fact is approved or a better solution is achieved by analyzing the available information that was provided by emitters [9]. The points of divergence within educational dialogues depict moments where creativity appears, because new ideas are introduced by emitters, ideas that can offer a new perspective which can lead to a new path and make to take into considerations other aspects [9]. It is not considered that the points of divergence emphasize moments where one of the participants are wrong, because the truth is not just only one view but a combination of views (convergent or divergent), it may appear as a result of interanimation [9]. Also, the points of divergence can be used as a measure of compatibility between students. When two students have several moments of disagreement along the chat discussion, it means that they have very different views, and each one of them can be a factor of horizon broadening for the other.

In terms of knowledge construction, convergences and divergences are different. Initially, in the chats we analyzed [4] each student is focused on his topic and generates a thread of discussion. The convergences of views suggest that links are generated between these threads, or can be seen as an approval on these discussion threads. The divergence of views suggests that the participants are focused on their own generated thread by being in contrariety with the other threads.

The purpose of the chat is to make its participants to engage in a discussion where they can share their personal knowledge and build new understanding of specific topics using the information shared by the other participants. The exchange of information is realized in more steps. The first step consists in bringing by the student the information that he thinks is appropriate to highlight the advantages of his topic. In the second step, the points of convergence and divergence will arise, because the other students will accept or refuse the ideas explored in the first step.

The approval among participants and the degree of positive interaction in the knowledge transfer are identified by finding the points of convergence and divergence between their views. The environment in which the discussion is taking place should be an encouraging one, where students can express themselves freely, so disagreements should not be made in such a way that a conflict may start or may inhibit other students in sharing new ideas. Disagreements should be a way in which the diversity of information must be

expanded, in other words, the new ideas shared may help others to rectify their beliefs or learn something new.

## IMPLEMENTATION

In this section, it is explained the system and the way in which the points of convergence and divergence are detected in chats. These chats have a special structure, which is different from the structure of chats that take place in social networks. This structure is represented in XML format and consists in: nickname of the utterer, the unique id of the utterance emitted by the utterer, the unique id of the explicitly referenced utterance (or -1 if there is not such a link was provided by the user), the time when the utterance is entered, and the text of the utterance [4]. The referenced utterance means that the actual utterance is a response to the referenced one, the echo of the previous one affects the actual one.

In Figure 1 are showed the main steps that the system performs.

### Step 1

*Chat parsing and explicit links detection* - in this phase the chat is parsed from the XML format, and the information contained in that format is saved in data structures that will facilitate the processing. Also, in this phase it is realized the detection of the explicit links, that are pointed by the chat participants when entering a new utterance. An explicit link is a link between two utterances, where usually the later utterance is in response to the earlier one.

### Step 2

*Detection of repetition links* – in this phase all utterances are divided into groups of a given length (ex: 20, 40, 100). The higher the length of a group is, the fewer groups will result after division. For every group are found the most frequent words that are used by participants using the NLTK library (which allows to access natural language processings such as stop words removing, tokenization, and lemmatization [5]). Unfortunately, most of the frequent words are not relevant, and will be of no help in finding agreements and disagreements. Because of that, after the previous step, a window with all of the most frequent words for every group is displayed, where the user can choose the most relevant words for every group.

Then, repetition links are identified, where a repetition link is a link between two utterances from the same group and in the text of the utterances is found the same word from those selected in the previous step.

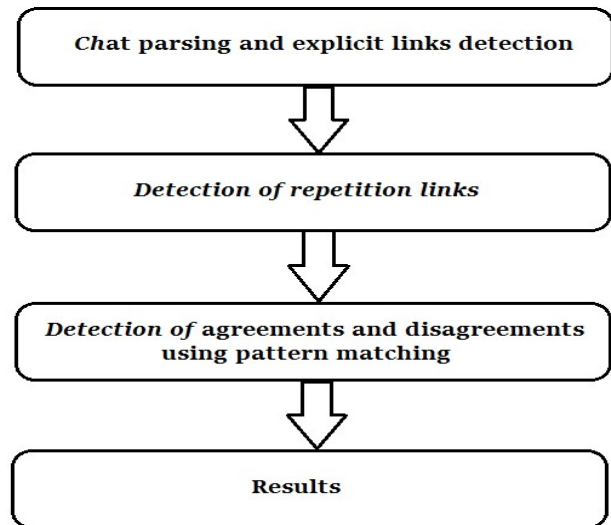


Figure 1. Implementation steps of the system

### Step 3

*Detection of agreements and disagreements using pattern matching* - in this phase a pattern matching module is used to detect points of convergence and divergence in chat. Explicit links are between two utterances, and to detect an agreement or disagreement in these links, the pattern matching module (see below) is used on the utterance with higher id (this utterance is considered a response to the utterance with lower id) and see if matches on any regular expression.

### Step 4

*Results* – in this phase are shown in text format the points of convergence and divergence and a graphic representation of how this moments can be interpreted

### The Pattern Matching Module

The pattern matching module, inspired from that one developed in the project LTfLL [10], can be used for finding occurrences in a chat that match a given expression. The module receives as input the text of the utterance and the expression that should be matched. The output is the matching solution if it exists.

This tool is suitable and very helpful for achieving the purpose of this task because in inter-human communication agreements and disagreements are expressed by specific patterns. These patterns will be the matching expressions used by the pattern searching module that will find the moments of interests in a chat, discussed in this paper.



In Table 1 are presented the elementary expressions that can be used to create complex expressions that the pattern matcher can use to detect points of convergence and divergence in utterances.

Elementary expressions	Matching
word	any occurrence of this word
„text”	any occurrence of this text (more words)
*	any string (even null string)
<D „stem”>	any word which has the stem „stem”
<S „word”>	any word that is a synonym of the word „word”
<NN>, <PRP>, <VBD>	matches any word that was annotated by the part of speech tagger with that label

Table 1. Elementary expressions

Operators help to create more complex expressions by combining more elementary expressions. Examples are shown in Table 2 and there is also presented their meaning and the way in which these operators can be used.

Operator	Symbol	Matching of new expression
Concatenation	Exp1 Exp2	The resulting expression matches a text if there is text1 that matches Exp1 and text2 that matches Exp2 and text = text1 + whitespaces + text2.
AND	Exp1 & Exp2	The resulting expressions matches text only if Exp1 matches text and Exp2 matches text.
OR	Exp1   Exp2	The resulting expressions matches text only if Exp1 matches text or Exp2 matches text.

Table 2. Pattern matcher’s operators for creating complex expressions

For grouping the expressions normal parenthesis can be used. By default, the evaluation of expressions is made by the pattern searching module from left to right. For example, Exp1 Exp2 & Exp3 is interpreted as (Exp1 Exp2) & Exp3 which is far different from Exp1 (Exp2 & Exp3).

In Table 3 are shown some expressions that the pattern searching module is using to detect the moments of interest from chat.

## RESULTS

The analysis of points of convergences and divergences were made on a chat which is in xml format. In the first part of this section will be shown the Convergences/Divergences between participants. In Figure 2 are shown this relations between participants represented graphically.

Expressions for detecting agreements		Expressions for detecting disagreements	
Expression	Examples of matches	Expression	Examples of matches
So (<VB>   <VBP>   <MD>) <PRP>	So do I	(<PRP> (<VB>   <VBP>   <MD>) not)   (<PRP> <VBP> <VB>)	I don’t think so I do not I do not think so
<PRP> too	me, too	<PRP> (<VB>   <VBP>   <MD>) either	I don’t either! I can’t either! I won’t either! I didn’t either! I wouldn’t either!
Neither (<VB>   <VBP>   <MD>) <PRP>	neither do I	I don’t agree	I don’t agree
(<DT>   <PRP>) (<VBZ>   <VB>) true	that is true it is true	Not at all	Not at all
<PRP> (<MD>   <VBP>   <VBD>) too	I will too, I can too, I did too, etc.	I don’t agree	I don’t agree

Table 3. Expressions for detecting points of convergence and divergence

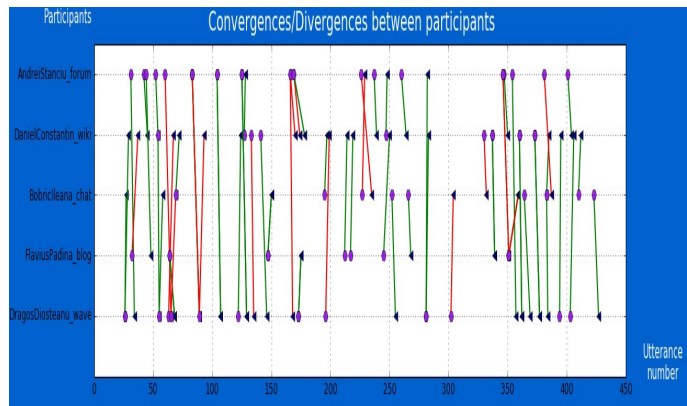


Figure 2. Convergences (in green) and Divergences (in red) between participants

Because there are too many moments to analyze, for exemplification we will focus on what is identified by the system between utterances 150 and 200. In Figure 2 is shown that in this interval were found seven moments of interest, where four of them are divergent (marked with red line) and three of them are convergent (marked with green line). From the graph is hard to detect the exact utterances that are implied in a given moment. To identify the utterances more precisely, a textual representation is provided, which shows that the first moment identified after the utterance 150 begins at utterance 166.

“DIV Matched expression: <PRP> (<VB> | <VBP> | <MD>) wrong  
 The referred utterance:  
 User: AndreiStanciu\_forum  
 ID: 166  
 Text: wikis, by definition, are made for public wide use  
 Utterance which express disagreement:  
 User: DragosDioșteanu\_wave  
 ID:168  
 Text: You are wrong here, Costa can tell us more”

For the next moments is hard to identify on the graph which one was first. From the textual representation, the moments are extracted :

“DIV Matched expression: Not necessarily  
 The referred utterance:  
 User: AndreiStanciu\_forum  
 ID: 166  
 Text: wikis, by definition, are made for public wide use  
 Utterance which express disagreement:  
 User: DanielConstantin\_wiki  
 ID:170  
 Text: not necessarily. we can restrict access to certain articles. that way” our wiki will have some sections for the general public who want to find out some details about our technologies

“CONV Matched expression: agree  
 The referred utterance:  
 User: AndreiStanciu\_forum  
 ID: 169  
 Text: yes, they have wikis that are publicly available, with public information, for the everyday user that takes an interest in that company's products  
 Utterance which express agreement:  
 User: DanielConstantin\_wiki  
 ID:178  
 Text: We can include that information in a wiki too. but I agree with Dragos's suggestion. Besides, the main purpose of our wiki would be to include technical details”

“DIV Matched expression: No way  
 The referred utterance:  
 User: AndreiStanciu\_forum  
 ID: 169  
 Text: yes, they have wikis that are publicly available, with public information, for the everyday user that takes an interest in that company's products  
 Utterance which express disagreement:  
 User: DanielConstantin\_wiki  
 ID:174  
 Text: besides there is no way that you can structure a database to make information easier to access than a wiki”

“CONV Matched expression: agree with <PRP>  
 The referred utterance:  
 User: DragosDioșteanu\_wave  
 ID: 173  
 Text: We can use a web site or a blog for doing that Andrei.  
 Utterance which express agreement:  
 User: FlaviusPadina\_blog  
 ID:175  
 Text: i agree with you and we also can post news or updates for our products”

From those textual representations and the graphic representation from Figure 2, another moment of divergence

was found between utterances (166, 170), and two moments of convergence were found between utterances (169, 178) and (173, 175). Also, another moment was identified between utterances (169, 174) which doesn't really express a disagreement, but is considered a point where new information is shared with the other participants.

## CONCLUSIONS

The proposed system aims to implement a method that is able to analyze chats in xml format and to detect points of convergence and divergence. The accuracy of the results depends on the agreement and disagreement patterns that are selected which must be as precise as possible. The accuracy of the system can be increased if more patterns for detection agreements and disagreements are added. Another way to improve the results is to use better patterns that are suitable to detect points of convergence and divergence and no other matchings that would not be considered adequate. This system is useful because makes possible to find the degree in which the students participate to the process of knowledge building.

**Acknowledgments.** This research was partially supported by the FP7 2008-212578 LTFLL project.

## REFERENCES

1. Trăușan-Matu, S., Chiru, C., Bogdan, R., Identificarea actelor de vorbire in dialogurile purtate pe chat, in Ștefan Trăușan-Matu, Costin Pribeanu (Eds.) *Interacțiune Om-Calculator*, Editura Printech, București, (2004) pp. 206-214
2. Stahl, G. *Group cognition: Computer support for building collaborative knowledge*. Cambridge, MA: MIT Press, (2006).
3. Trăușan-Matu, S., Dascalu, M., Rebedea, T., PolyCAFe—automatic support for the polyphonic analysis of CSCL chats, *International Journal of Computer-Supported Collaborative Learning*, Volume 9(2), Springer, (2014), pp. 127-156
4. Trăușan-Matu, S., Dascalu, M., Rebedea, T., Gartner, A., *Corpus de conversații multi-participant și editor pentru adnotarea lui*, Revista Română de Interacțiune Om-Calculator 3(1), (2010), pp. 53-64
5. <http://www.nltk.org/> (last accessed 30<sup>th</sup> June 2017)
6. <https://ceasefiremagazine.co.uk/in-theory-bakhtin-1/> (last accessed 30<sup>th</sup> June 2017)
7. Bakhtin, M.M. *The dialogic imagination: Four essays* (trans: Emerson, C. & Holquist, M.). Austin and London: The University of Texas Press, (1981)
8. Trausan-Matu, S., *The Polyphonic Model of Hybrid and Collaborative Learning*, In Wang, F.,L., Fong, J., Kwan, R.C., *Handbook of Research on Hybrid Learning*

*Models: Advanced Tools, Technologies, and Applications*, Information Science Publishing, Hershey, New York, (2010), pp 466-486

9. Trausan-Matu, S., Stahl, G., Sarmiento, J., Supporting Polyphonic Collaborative Learning, E-service Journal, vol. 6, nr. 1, Indiana University Press, (2007), pp. 58-74.
10. Trausan-Matu, Stefan; Dessus, Philippe; Rebedea, Traian; Mandin, Sonia; Villiot-Leclercq, Emmanuelle; Dascalu, Mihai; Gartner, Alexandru; Chiru, Costin; Banica, Dan; Mihaila, Dan; Lemaire, Benoît; Zampa, Virginie; Graziani, Eugène, LTfLL - D5.2 Learning support and feedback, (2009)

# Gesture-based Visual Analytics in Virtual Reality

**Mihai Pop**

Technical University of Cluj-  
Napoca

str. G. Barițiu 28, 400027,  
Cluj-Napoca, România  
mihai.m.m.pop@gmail.com

**Adrian Sabou**

Technical University of Cluj-  
Napoca

str. G. Barițiu 28, 400027,  
Cluj-Napoca, România  
adrian.sabou@cs.utcluj.ro

## ABSTRACT

This paper presents an approach to dynamic data visualization and manipulation using virtual reality as a means of display and the Leap Motion controller in order to interact with the virtual scene. The data and related APIs provided to the virtual reality application are based on a separate web application which provides REST APIs in order to get and further manipulate existent data. The algorithm pipeline is managed by using the Unity Game Engine, which provides a means to express control and logic over the sensory inputs of the Leap Motion controller. The application is built by using the Unity Game Engine and it is streamed on any Android phone through RiftCat Vridge software solution for streaming application over Wi-Fi, while interpreting the phone's sensor input and translating it into virtual head movements. The issues from the standpoint of a corporate need for a more interactive and natural experience in viewing and interacting with big data.

## Author Keywords

Visual Analytics; Virtual Reality; Leap Motion; Gestures detection; RiftCat Vridge; Node server application;

## ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

## General Terms

Virtual Reality; Data analytics.

## INTRODUCTION

Over the last couple of years, the virtual reality scene has become more and more popular. Part of the reason is that computational power has become less expensive and readily available to those who study and create projects with such needs. From its early cumbersome beginnings, Virtual Reality has become widely spread in the game industry through simulation games using handheld controllers and a head-mount for displaying the virtual scene. Though this has been the main use case for virtual reality applications, new use cases are emerging lately from the corporate world.

When compared to the conventional means of analyzing big data, virtual reality data analytics is more appealing and visually pleasing than scrolling through endless table rows. There are large datasets which need to be sorted, viewed, compared, enhanced. A virtual reality solution would offer the user a means of interacting with the data in a more natural manner; through gestures, and physical user interface objects which mimic the behavior of real-life buttons, sliders and knobs.

As a response, this need of a more convenient and natural way to approach data analytics has fueled the development of an application designed to receive large sets of data and display and manipulate them in order to fulfill the user's data queries and study.

The context for this virtual reality application is set by the corporate need of a better data analytics visualization and manipulation approach. Having this in mind, an auxiliary server-side application has been developed in order to substantiate real-world server behavior that we plug into our analytics application. The server application provides a basic REST API for querying and filtering data.

Then, the users queries data by requesting it with the appropriate request URL. The requesting of data can be done through the use of hand gestures in front of the Leap Motion controller. For testing purpose, the data is currently represented as a 3D array of cubes. This array of cubes is generated at runtime through an instantiation algorithm which takes into account all data from the server application and maps it accordingly to Unity game objects with attached data storing scripts for the server-side data.

So, by mapping the server-side data into compatible Unity game objects, we can apply our desired logic onto the virtual scene providing the user many features for an interactive and immersive experience in order to suit his goals through simple hand gestures.

The proposed solution makes use of the phone's gyroscopic information in order to track head movements and to supply this information to the cameras in the virtual scene, allowing them to rotate and translate with respect to the tracked head movements. Also, we use the Leap motion controller for its hand tracking and development library which provides a toolkit for defining custom gestures to be

detected and many user interface functionalities which make the application user friendly and smooth.

The chosen solution uses custom gestures built upon predefined elements for certain actions which are all described as Unity classes or logic gates between several gesture detectors from the Leap Motion Controller library.

This paper elaborates on a dual application solution which tries to solve the issue of big data manipulation using natural gestures in order to achieve this goal.

## RELATED WORK

A variety of techniques for virtual reality interaction already exist, due to the high demand in the gaming market for such applications and games. As a result, several iterations of relevant libraries and APIs have been created to suffice the development of virtual reality applications. As virtual reality further matures, it will only get easier to start developing and creating an interactive virtual reality application due to the advances of previous developers and the documentation of their features.

Ridder et al. [1] used Virtual Reality and Augmented reality coupled with gestures to create an immersive environment for visualizing fMRI data. fMRI or Functional magnetic resonance imaging consists of a functional neuroimaging procedure using MRI technology in order to measure brain activity by detecting changes related with blood flow. They suggested that VR/AR can potentially allow a reduction in visual clutter and lets users keep their focus on visualizations by allowing them to navigate the data abstractions in a natural way.

Moran et al. [2] used interactive Virtual Reality to manipulate Twitter datasets and to visualize them in the original geospatial domain. By using emerging technologies, they created a fully immersive tool that promotes visualization and interaction and that can help ease the process of understanding and representing big data.

Olshannikova et al. [3] provide a multidisciplinary overview of the research issues and achievements in the field of Big Data and its visualization techniques and tools. They discuss the impacts of new technologies, such as Virtual Reality displays and Augmented Reality helmets on the Big Data visualization as well as to the classification of the main challenges of integrating the technology.

Hackathorn and Margolis [4] outline the objectives for analytical reasoning and immersive data spaces, followed by suggestions for the design and architecture of data worlds. Finally, they describe current work for building data worlds.

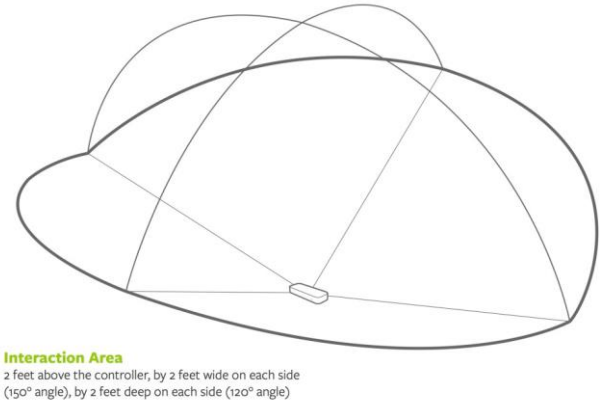
## THE LEAP MOTION CONTROLLER

The Leap Motion Controller device [6] consists of two cameras and three infrared LEDs. These track the infrared

light with a wavelength of 850 nanometers, which is of course outside the visible light spectrum.

Thanks to its wide-angle lenses, the device has a large interaction space of 0.226 cubic meters, which takes the shape of an inverted pyramid – the intersection of the binocular cameras' fields of view (Figure 1).

With the Orion beta software (the current iteration of the software development kit for the Leap Motion device), this has been expanded to 2.6 feet (80 cm). This range is limited by LED light propagation through space, since it becomes much harder to infer your hand's position in 3D beyond a



**Figure 1. Leap motion interaction area [5]**

certain distance.

At this point, the device's USB controller reads the sensor data into its own local memory and performs any necessary resolution adjustments. This data is then streamed via USB to the Leap Motion tracking software.

The data takes the form of a grayscale stereo image of the near-infrared light spectrum, separated into the left and right cameras. Typically, the only objects you'll see are those directly illuminated by the Leap Motion Controller's LEDs. However, incandescent light bulbs, halogens, and daylight will also light up the scene in infrared. You might also notice that certain things, like cotton shirts, can appear white even though they are dark in the visible spectrum.

Once the image data is streamed to your computer, it is time for complex mathematical computations. Despite popular misconceptions, the Leap Motion Controller doesn't generate a depth map – instead it applies advanced algorithms to the raw sensor data.

The Leap Motion Service is the software on your computer that processes the images. After compensating for background objects (such as heads) and ambient environmental lighting, the images are analyzed to reconstruct a 3D representation of what the device sees.



Next, the tracking layer matches the data to extract tracking information such as fingers and tools. Our tracking algorithms interpret the 3D data and infer the positions of occluded objects. Filtering techniques are applied to ensure smooth temporal coherence of the data. The Leap Motion Service then feeds the results – expressed as a series of frames, or snapshots, containing all the tracking data – into a transport protocol.

### NODE REST APPLICATION WITH MONGODB DATABASE

As mentioned in the introduction, there is an auxiliary application which interacts with the virtual reality application, providing it with the dynamic data and advanced query capabilities. In the following, there is a brief description of the core concepts of such an application.

One of the most used approaches is Representational State Transfer – REST – because it is an open approach for lots of conventions that are used for consumers of your API. The way this transfer is made is determined by the resources provided by your API.

RESTful applications use HTTP requests to perform four operations termed as CRUD (C: create, R: read, U: update, and D: delete). Create and/or update is used to post data, get for reading/listing data, and delete to remove data.

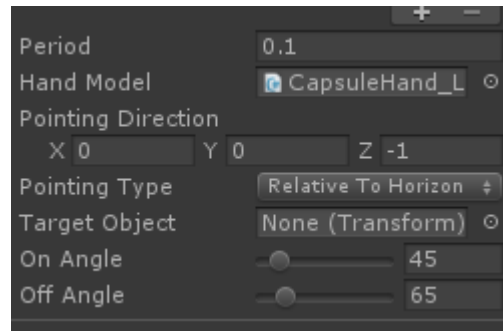
MongoDB [7] is an open source database that uses a document-oriented data model. MongoDB is one of several database types to arise in the mid-2000s under the NoSQL banner. Instead of using tables and rows as in relational databases, MongoDB is built on an architecture of collections and documents. Documents comprise of sets key-value pairs and are the basic unit of data in MongoDB. Collections contain sets of documents and function as the equivalent of relational database tables.

So, the server-side application is a REST application which uses Node [8] for server emulation, a server-side solution for JavaScript, and in particular, for receiving and responding to HTTP requests. The application is written in JavaScript language, for a clean and rapid development.

### LEAP MOTION GESTURES

The Leap motion software development library by itself does not contain definition for complex gestures [9]. The Leap motion API contains definitions for the mapped human body parts (Hand, Arm, Bone, Finger, etc.) and also definitions for concepts like Frame (the clipped viewport in which a Leap object exists) or Controller (you can access the actual controller attributes and methods) or other non-physical control concepts.

So, in order to define more complex gestures, we can make use of the Leap Motion Detection Utilities Module. The detection utilities are a set of scripts in the core asset package that provide a convenient way to detect what a



**Figure 2. The configuration of the Extended Finger Detector**

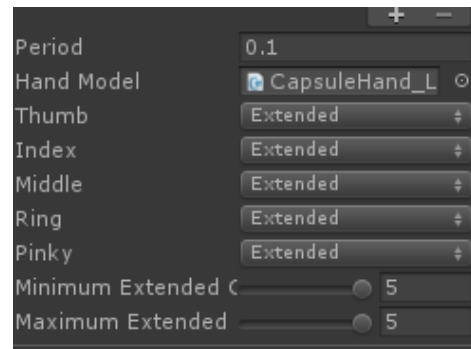
user’s hand is doing. For example, you can detect when the fingers of a hand are curled or extended, whether a finger or palm are pointing in a particular direction, or whether the hand or fingertip are close to one of a set of target objects.

Detectors can be combined in order to declare complex gestures. This can be done by using a Logic Gate. The Detector Logic Gate provided by this module is in itself a detector which combines two or more detectors to determine its own state.

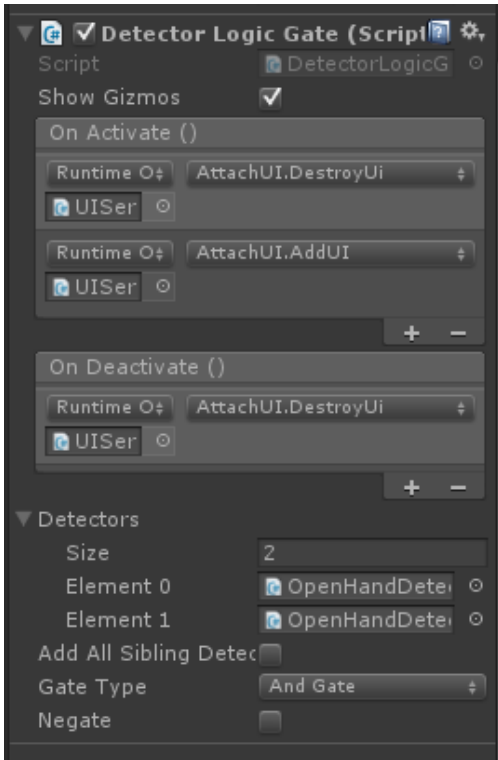
Detectors ultimately dispatch standard Unity events upon activation or deactivation. This provides a convenient means to hook different function handlers.

For instance, when we would want to detect if the user has his hand camera-facing and open we would assemble the following detectors:

1. Extended Finger Detector (Figure 2) - configure the component so that all fingers must be extended.
2. Palm Direction Detector (Figure 3)
  - a. Pointing direction = (0, 0, -1)
  - b. Pointing type = Relative to Horizon
  - c. On and off angles: set as desired



**Figure 3. The configuration of the Palm Direction Detector**

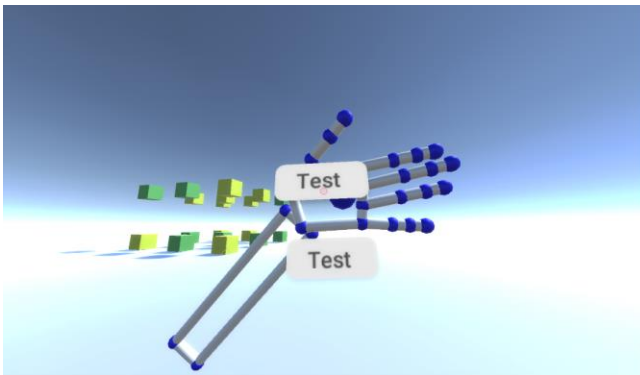


**Figure 4. Detector Logic Gate configuration**

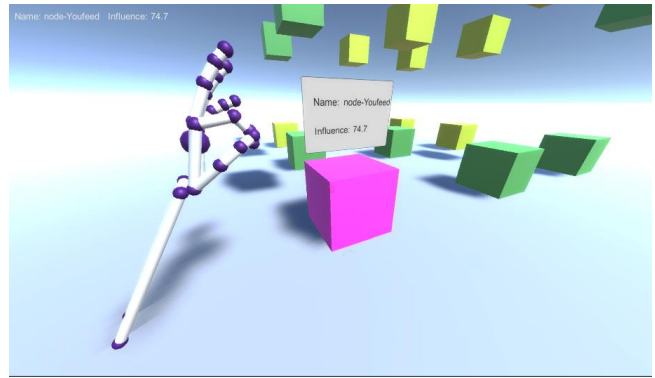
After that, we can link these two detectors in the Detector Logic Gate and link other C# scripts and functions to further our desired logic or use case (Figure 4).

This minimal setup can facilitate powerful results. The above example links the detection of a camera-facing open hand to a function call from a separate file, which provides an UI attachment to the left palm of the user (Figure 5).

Another such gesture is pinching which is defined similarly through the use of detectors. The effects of pinching with the left hand while looking at a given cube will cause it to highlight and display a panel over it with its mapped data



**Figure 5. The instantiation of UI components on camera-facing open hand gesture**



**Figure 6. Pinching with left hand highlights cube and displays its information**

from the server-side application (Figure 6).

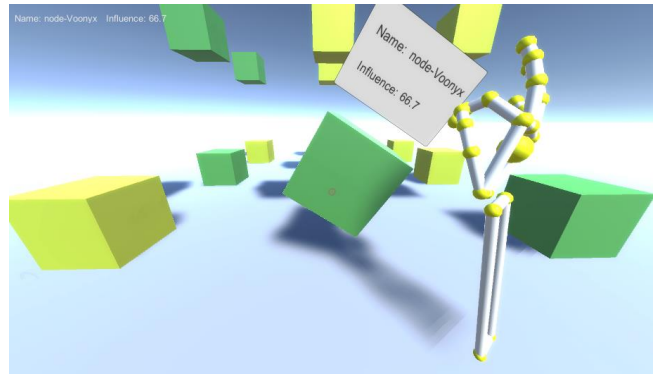
The data from the server-side application is parsed before attaching it in the form of data scripts to Unity game objects. The parsed JSON data is further integrated in the application by being mapped as a custom list of ResourceNode objects, which are a custom class linked to the project namespace.

If the user would pinch with his right hand while looking at any certain cube, that cube will change position in the Unity scene hierarchy and shall be moved under the RTS node which stands for Rotate Translate Scale node (Figure 7).

This Node has scripts attached to it which rotate, translate and scale any mounted game object. It also has capabilities for one and two-handed rotation.

### SERVER APPLICATION IN DEPTH

The server application uses a MongoDB database on port 27017, the standard MongoDB port, this can obviously be changed if conflicting settings. The application runs on port 3000 and has different URL mappings, in order to provide consumption to its API.



**Figure 7. Pinching with right hand mounts the looked at cube in the RTS node**



**Figure 8. Main get URL for getting all database entries**

Of course, in order to start the MongoDB local service on port 27017 we would need to execute this statement in a terminal: `mongod --dbpath "C:\data"`. This statement just starts the “mongod” process having as database path, the above-mentioned path.

The server mappings are a collection of URLs for getting information, which support search parameters. Examples of the URL mappings are as follows:

1. <http://localhost:3000/api/data>  
This is the main GET URL, which gets all data in the database (Figure 8)
2. <http://localhost:3000/api/data?name=Youfeed>  
This gets the node with the name Youfeed.
3. <http://localhost:3000/api/data?influenceGt=20&influenceLt=40>  
This gets the nodes with the influence key between 20 and 40 (Figure 9).

These filtering operations can essentially be endless because MongoDB offers rich functionality for collection manipulation.

### Server Data Mapping

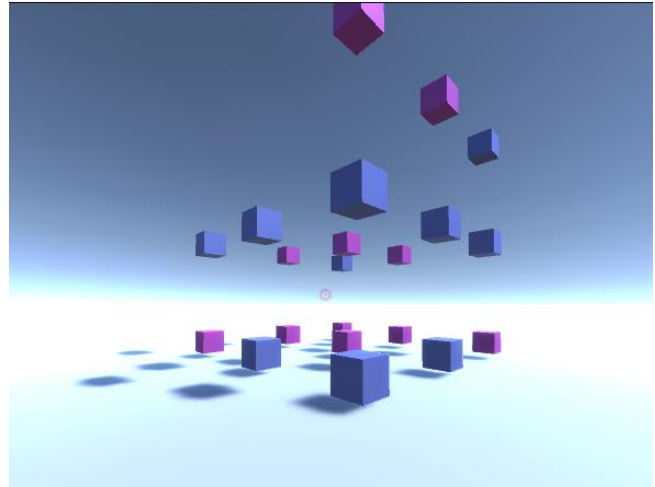
In order to represent the server-side data as cubes, there must be defined some transitory objects between the primitive cube game objects and the actual raw data.

The typical pipeline of processing server data in our application is:

1. Execute Start void function from Instantiation Service script

This function calls another function which does the actual request.

2. Check if erroneous response from the function which made the server call. If everything is ok, we call a function which builds the data



**Figure 9. Nodes between influence range [20, 30]**

3. The function which builds the data does the parsing of the JSON server response. Each value is parsed as the appropriate data type it should have. A list of Resource Nodes is created, which aid in mapping the JSON data. Resource Node is a custom class which belongs to the project namespace After parsing, call a function which creates nodes.
4. The function which creates the nodes loops over the list of Resource Nodes and creates new Cube object which is also a custom class that has attributes for parsed values and methods for setting them and also a render method which adds a game object to the scene with mapped properties
5. The position of each cube is computed by the place cubes function which basically tries to construct a cube of cubes

## IMPLEMENTATION

### The Unity game engine

For the implementation, there was an extensive research as to how to assure compatibility the many components at play in this project. The final versions chosen for the components are:

1. Leap Motion SDK 3.1.2
2. Leap Motion Core SDK 4.1.4
3. Leap Motion Attachments Module 1.0.4
4. Leap Motion UI input module 1.0.0
5. Unity Game Engine 5.3.4f1
6. RiftCat Vridge 1.3.3
7. Oculus Runtime 0.8

The reasoning for choosing these compatible versions was in order to stream the application on any android phone and make use of its sensors by not faking sensor input and using the actual phone's sensor information.

## CONCLUSION

In the world of virtual reality, technology is advancing drastically and new implementations and solutions to problems we do not know we have arrive every day. The rise of the virtual reality world has fueled other industries to peer down and indulge into the advantages this industry brings.

Therefore, the use cases associated with virtual reality got more complex and more 'big players' got into the virtual reality game. In consequence, there is a fueled need in the corporate world for a solution to common reporting problems through the means of virtual reality.

Trying to provide a user friendly and straightforward solution, this application could be the start of the modern way to visualize and manipulate data analytics by using virtual reality.

This paper presented a dual application solution in order to give a user the opportunity to do his queries and tasks through a virtual interface in which everything is controlled by gestures and intuitive design.

Future work includes a better user experience and integrating with email and text-to-speech APIs.

## REFERENCES

1. M. de Ridder, Y. Jung, R. Huang, J. Kim and D. D. Feng. Exploration of Virtual and Augmented Reality for Visual Analytics and 3D Volume Rendering of Functional

Magnetic Resonance Imaging (fMRI) Data. *2015 Big Data Visual Analytics (BDVA)*, TAS (2015), 1-8

2. A. Moran, V. Gadepally, M. Hubbell and J. Kepner. Improving Big Data visual analytics with interactive virtual reality. *2015 IEEE High Performance Extreme Computing Conference (HPEC)*, (2015), 1-6.

3. Ekaterina Olshannikova, Aleksandr Ometov, Yevgeni Koucheryavy and Thomas Olsson. Visualizing Big Data with augmented and virtual reality: challenges and research agenda. *Journal of Big Data* 2, 1 (2015), 22.

4. R. Hackathorn and T. Margolis. Immersive analytics: Building virtual data worlds for collaborative decision support. *2016 Workshop on Immersive Analytics (IA)*, (2016), 44-47.

5. Alex Colgan. How Does the Leap Motion Controller Work?. Leap Motion Blog. (2014).

<http://blog.leapmotion.com/hardware-to-software-how-does-the-leap-motion-controller-work/>

6. Leap Motion Inc. Leap Motion. (2017).

<https://www.leapmotion.com/>

7. MongoDB Inc. MongoDB. (2017)

<https://www.mongodb.com/>

8. Node.js@ JavaScript runtime (2017)

<https://nodejs.org/en>

9. Leap Motion Inc. Using the Tracking API. Leap Motion Developer (2017).

[https://developer.leapmotion.com/documentation/python/devguide/Leap\\_Guides2.html](https://developer.leapmotion.com/documentation/python/devguide/Leap_Guides2.html)

# Considerations for a Virtual Training System to Improve Job Interview Skills for Software Engineers

**Iulia-Cristina Stănică**  
Department of  
Engineering in Foreign  
Language, University  
Politehnica of  
Bucharest  
Splaiul  
Independentei 313,  
Bucharest, Romania  
iulia.stanica@gmail.  
com

**Florica Moldoveanu**  
Computer Science and  
Engineering  
Department,  
University  
POLITEHNICA of  
Bucharest  
Splaiul  
Independentei 313,  
Bucharest, Romania  
florica.moldoveanu@cs.  
pub.ro

**Maria-Iuliana  
Dascălu**  
Department of  
Engineering in Foreign  
Language, University  
Politehnica of  
Bucharest  
Splaiul  
Independentei 313,  
Bucharest, Romania  
maria.dascalu@upb.ro

**Alin Moldoveanu**  
Computer Science and  
Engineering  
Department,  
University  
POLITEHNICA of  
Bucharest  
Splaiul  
Independentei 313,  
Bucharest, Romania  
alin.moldoveanu@cs.p  
ub.ro

**Dan Alexandru  
Mitrea**  
Department of  
Engineering in Foreign  
Language, University  
Politehnica of  
Bucharest  
Splaiul  
Independentei 313,  
Bucharest, Romania  
danalexmitrea@gmail.  
com

## ABSTRACT

Nowadays, engineers have to face a lot of challenges before going to a job interview. Introversion, insecurity or lack of experience are some of the reasons that can lead to poor interview performances. Through training, such difficulties can be overcome, leading to significant skill improvement. Our paper presents the advantages of using emerging technologies to facilitate job interview training. Our proposed model for a training system is based on virtual reality and chatbots. Hard skills and soft skills training, various 3D virtual environments and domains modularity could ensure the personalization of the system for any job training situation.

## Author Keywords

Interview; engineering; training; chatbot; virtual reality.

## ACM Classification Keywords

H.5.1. Information interfaces and presentation (e.g., HCI):  
Multimedia Information Systems.

## General Terms

Human Factors; Documentation; Theory.

## INTRODUCTION

In the 21<sup>st</sup> century, engineering is one of the most appreciated and successful domains. Engineers must have various competencies, including mathematics, science, economics, or social knowledge in order to create, research and innovate. The field is quite large, as engineers are capable of creating impressive inventions related to mechanics, chemistry, electricity or software.

It is thus obvious why small, medium or big companies are always searching for competent engineers. Over the last few years, a small tendency of growth was observed regarding the employability of engineers on the job market, and it is most likely that this increase will continue in the following years. Software engineering and IT fields are

probably the most successful ones, having the biggest employment growth after the economic recovery [1].

Even if the numbers show that software engineers are sought in all companies, their interview performances do not always have the expected results. Engineers usually have to face numerous difficulties when applying for their first job. Even worse, some of these problems still exist after the first interview: introversion, insecurity, lack of experience or lack of social contact are just a few reasons which can lead to poor interview performances.

There are many articles that try to give pertinent advice for engineers in order to help them prepare for an interview. There are various aspects which must be carefully arranged, as they are essential for performing well at an interview [2]:

- Revising previously acquired knowledge from university;
- Getting ready to handle unpredictable scenarios;
- Analyzing important personal accomplishments related to engineering;
- Researching company-related information and issues;
- Remaining calm and true to oneself.

Unfortunately, all these aspects are difficult to accomplish by young graduates, with little to no experience. Therefore, we can say that the real key for obtaining good performances at a job interview is training. By definition, training is an activity which must be undertaken by a person in order to achieve a certain degree of knowledge or competencies. By training, new information and instructions are communicated to the trainee, helping them improve and gain the required skills [3]. If we practice some standard questions and revise our previous knowledge, our chances of obtaining better results at interviews increase. Learning through experience is always better, as the Chinese philosopher Lao Tse says: “If you tell me, I will listen. If you show me, I will see. But if you let



me experience, I will learn”. Training methods relying on theory are not always successful, and job interview training should be based on practice.

There are several methods of simulating an interview situation: practicing in front of a mirror, asking a friend for help or, even better, seeking advice from career counseling experts. Each one of these possibilities has some advantages and some drawbacks. Firstly, the idea of practicing in front of a mirror can boost one’s confidence, help them put their ideas in order and formulate a possible speech. On the other hand, since there is no actual feedback, the person cannot be sure if their speech is a good one; they are not taking part in a dialogue and can only practice the questions they perceive as being relevant. If they practice with a friend, there is some feedback, but the lack of technical (or even social) experience can make the interlocutor formulate irrelevant questions. Finally, the idea of hiring a career counseling expert can offer the most accurate experience. Unfortunately, several counseling sessions are needed, which, in turn, can be quite expensive. Human subjectivity is a permanent factor, which can influence erroneously our training process

Our current paper will propose an innovative training system which can be used to facilitate job interviews for engineers. In the following chapters, we will present some emerging technologies which can be used to create training systems, we will analyze the existing computer-based systems for job interview training and finally, we will propose our model based on emergent technologies.

**EMERGING TECHNOLOGIES FOR TRAINING SYSTEMS**

We live in an era of technology, so the time has come to try and integrate it in as many domains as possible. Traditional training methods can therefore be replaced by the use of emerging technologies, which can thus facilitate the training process.

**Virtual Reality**

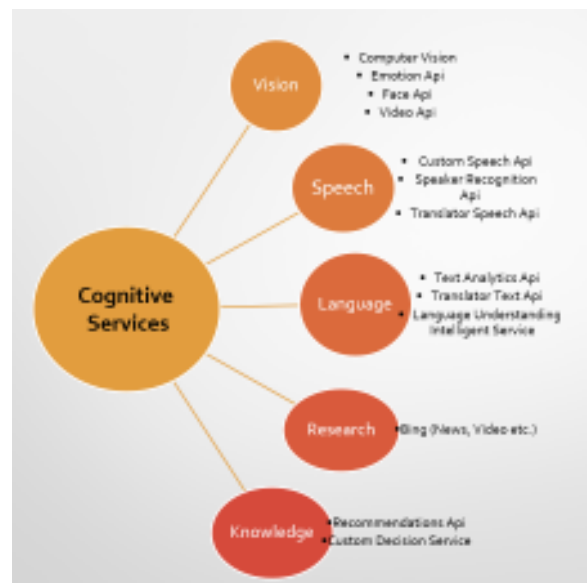
Such an example is Virtual Reality (VR), a current trend, an emerging technology which has the capacity of creating virtual environments with the help of computer-based means, simulating, thus, real-life situations. Since 1950, when the term was coined, this technology has never ceased to evolve. Additional devices, such as head mounted displays (HMD), joysticks, gloves or steering wheels are invented and perfected continuously, providing the authenticity of the immersion. In the past, we were talking about “classical virtual reality”, focusing on equipment. Now, we are talking about “immersion-centered VR”, which concentrates on stimulating the two main senses of the user (sight and hearing) in order to create an authentic, realistic and immersive experience [4] [5]. The capacity of creating a great variety of virtual environments, the accessibility and flexibility of the process, or the reduced costs and instant feedback are just a few reasons which justify the idea of using virtual reality to create training

systems. We have analyzed several researches regarding the use of VR for training and we have realized that it is a promising technology, which could be used in many domains, such as education [6], medicine [7], army [8], architecture [9] or entertainment [10].

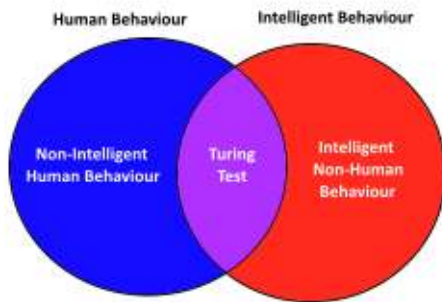
**Cognitive Services and Chatterbots Frameworks**

Microsoft’s cognitive services represent a group of APIs which use machine learning-based models in order to ensure a natural interaction between users and their various tools. By using cognitive services, algorithms from a large variety of fields can be exploited, such as Computer Vision, Voice Recognition, Recommender Systems, Face and Emotion Recognition, Text Analytics (Figure 1). They are based on artificial intelligence principles and can facilitate a great variety of tasks which require fast extraction of information from data. These APIs can be easily imported in any programming language and integrated into any platform, such as Android, iOS or Windows [11] [12].

A chatterbot or chatbot is a software-based program capable of simulating a conversation with human interlocutors, especially through the Internet [13]. The conversational partner might believe they are talking to a real human being, so chatterbots must be capable of passing the Turing Test. This test was developed by Alan Turing in 1950 and has the goal of verifying if a machine is capable of reproducing intelligent behavior equivalent to that of a human being. Passing the Turing Test can be challenging, as machines do not require just the simulation of intelligence: they should also be capable of reproducing unintelligent human behavior (Figure 2) [14]. Chatbot frameworks represent places where bots are actually built and where their behavior is conceived. There are many development frameworks for creating chatbots, most of them being based on natural language processing (NLP).



**Figure 1: Cognitive Services and their APIs.**



**Figure 2: Turing Test Principle.**

NLP applications take as input natural language and process it at low or high level in order to accomplish certain tasks (recognize parts of speech, find information or answer questions) [15]. Some well-known frameworks for chatbots are [16]:

- API.ai, which includes support for various messaging platforms, such as Facebook or Slack.
- Wit.ai, entirely free, ideal for software developers, capable of being integrated in many messaging APIs and SDKs.
- Microsoft Bot Framework, which offers the capacity of using multiple interconnected bots, able to interact in a natural manner with users, through text, email, Skype or Slack.

If we take into account all the previously analyzed aspects, chatterbots and virtual reality can represent useful tools for facilitating the training process, in various domains. For instance, sales departments can benefit from these emerging technologies as their time and budget limitations are the main obstacles preventing them from investing in their personnel's training. Regular training can involve high costs, can affect the company's productivity and does not include actual interaction with real clients. Therefore, many sales teams do not receive the appropriate training in order to achieve their maximum potential. A possible solution is the use of e-learning, but the experience is not immersive enough to ensure an accurate simulation of real-life situations. An alternative is represented by the use of virtual reality and chatbots – it can facilitate the transition of the training process from a static, often boring experience, to a dynamic, immersive one. The use of these emerging technologies can bring many advantages, such as the total implication and immersion of the trainee, reduced costs, great accessibility, flexibility and commodity of virtual learning. By using HMDs, employees who receive training can find themselves in the right position: face to face with the bot. This can ensure a profound, significant and relevant training experience [17].

Another situation presents the use of chatbots for security training in an enterprise. Two case studies conducted with 106 users and security specialists show qualitative progress

in the training process. By using an e-learning package, including a chatbot named Sally, various areas of questions, such as knowledge, behavior and attitude towards security, were tested. The attitude of trainees which participated to the case studies was more positive when chatterbots were used, as opposed to the traditional use of conventional e-learning for security training [18].

One of the domains which show great potential in the use of chatbots for training is human resources (HR). Even though we are still at the beginning of chatterbots' evolution, there are already several prototypes which bring many advantages in HR training. Chatbots can be an excellent alternative to the various tools employed in a company, which can be too complex or difficult to use. Chatbots are simple, intuitive and can be successfully used for HR functions and interacting with the employees. The use of this technology has many advantages, as chatterbots programs can be easily and rapidly implemented if there is already an internal communication system in the company. In addition, the recruitment process can also be facilitated by the use of chatbots, which are capable of filtering the spams received by the recruiters. Therefore, only the best candidates will get to interviews, making the recruiters' job easier [19]. Other advantages of using chatbots include the capacity of solving the problem of specific terminology and jargon for unfamiliar users or the objectivity and politeness of bots when having to deal with rude clients, intense traffic, repeated requests of the same information [20].

#### **COMPUTER-BASED SYSTEMS FOR JOB INTERVIEW TRAINING**

Job interview training can be accomplished through the use of online courses: Udemy, English Online Learning or Impact Factory are just a few examples of websites which offer video tutorials meant to train specific competencies. Recent technologies, such as virtual reality and chatbots can be combined in order to obtain better results.

Such an example is the program "Job interview with Molly Porter", developed by the universities of Yale and Northwestern in collaboration with professional rehabilitation specialists. The system is quite complex and allows users to practice computer-based job interviews in a stress-free environment. A practical simulation with a fictional interviewer (Molly Porter), user-based teaching material, interactive hiring process and constant, advanced feedback are some of the characteristics included in the system. Advanced simulation technologies make each conversation with Molly unique, helping the trainees acquire useful competencies and gain confidence during an interview. Feedback is given permanently by the virtual trainer during the interview and a complete critique is available at the end of the process. The Molly Porter system has also been tested with adults suffering from various mental disabilities, such as schizophrenia, autism or post-traumatic stress disorder. Results are encouraging, showing that similar systems could have a large specter of efficient

uses [21] [22]. Using virtual reality to train mentally disabled people is often the subject of current researches, their goal being that of increasing their employability and integration on the job market [23].

One essential aspect related to VR-based job interview training is the need of ensuring high immersion of the experience and a high degree of realism. Various studies analyze how realism can influence anxiety in a virtual environment and the efficiency of job interview training. One particular research conducted by specialists from the UK and South Korea shows the results of training in a virtual environment with various levels of realism. In order to create these different levels, the avatar changed from a realist photo to a cartoonish 3D model and finally, to a highly realistic 3D model. Results show that levels of stress were most elevated with the most realistic avatar, but the differences between situations were almost insignificant. This confirms the theory that anxiety is determined by the idea of taking part in a stressful situation, not by the graphical realism of the avatar [24]. The job interview situation might cause a lot of stress because it is outside the person's control, being also evaluative and competitive [25]. Therefore, the virtual environment should rather ensure the emotional immersion, by using head mounted displays and trying to reproduce the situation as accurate as possible (in terms of dialogue, user-bot interaction). VR-based systems are not meant to replace real interviewers, as there are specifically trained psychologists on that job which cannot be replicated by virtual reality. Their goal is to create a sufficiently immersive environment which can induce the appropriate emotional response from the user. Studies show that virtual environments for training job interview skills can be really successful, if an appropriate degree of immersion is provided [26].

**VR-JOB: AN INNOVATIVE MODEL FOR A JOB INTERVIEW TRAINING SYSTEM BASED ON EMERGING TECHNOLOGIES**

Our proposed system will mix together the advantages of both virtual reality and chatbots emerging technology in order to create a job interview training system. There is little research regarding the use of both virtual reality and chatterbots in order to improve job interview skills of software engineers. The two emerging technologies can bring many advantages when being combined, such as creating an immersive realistic environment, simulating accurately a great variety of job-training situations, providing knowledge from many technical fields, offering accurate real-time reactions to the trainee's responses. Already existing systems focus mostly on job interview training for social categories, which is significantly different from training for software engineers. "Job interview with Molly Porter", although complex and stable, focuses on work history, future plans and social skills, lacking the technical part, essential for software engineering interviews. Technical skills can be easily practiced and evaluated through a VR and chatterbot-based system, while

software engineers can show less adversity towards this genre of training. Furthermore, if we take into account how effective these systems are for training in various fields, we can be convinced that a well-designed system can prove to be successful for our purpose.

In the VR-JOB system, the chatbot is connected to a knowledge base, which is fed by a different component, managed by the administrator. The knowledge base can be either a database or a series of XML files which contain various information needed for training the chatbot: reactions to the candidate's answers, general facts, questions asked by the chatbot itself etc.

Firstly, the knowledge base will include information which is meant to test the "hard skills" of the candidate. The user will thus practice their technical knowledge, by receiving different questions from specific domains, related to the job description, such as data structures, algorithms, databases, web programming or various programming languages. For training "soft skills" such as social competencies, communication or personality related traits, the chatbot must be trained to ask general questions, related to hobbies, passions, extracurricular activities.

Artificial intelligence is a wide domain, therefore training a chatbot and ensuring its pertinent behavior can be a challenging task. We will thus limit the domain (the available inputs/outputs) in order to reduce the complexity of the system. The user must be capable of asking several questions (related to the company, the benefits of the job etc.), but the domain of these questions must be narrowed. Each candidate can create an account, login and select an enterprise from a predefined list where they want to apply for the job. Different virtual environments and different avatars will be generated based on the previously chosen enterprise. All the high-level functionalities are specified in the use case diagram (Figure 3).



**Figure 3: VR-Job Use Case Diagram.**

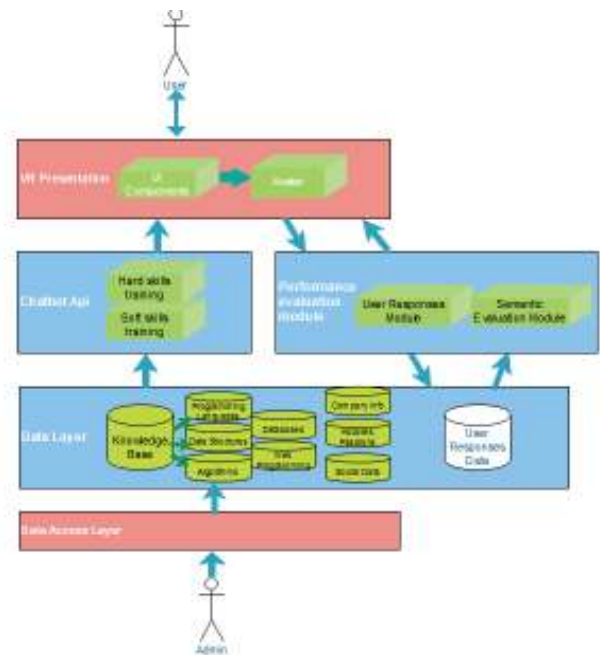
The evaluation of the candidate's performance is based on the answers given to the technical and soft skills questions. For the technical questions, which will be saved in the database / XML files, the system will check the presence of certain keywords. If they are present, the answer will be validated as being correct; if not, it will be erroneous. In order to achieve this, various semantic models could be used, based on ontologies [27] or NLP. The performance of this type of system is usually evaluated based on "semantic tasks", such as synonyms detection, analogy recognition or verb selection preferences. As far as soft skills are concerned, feedback can be given by combining semantic models and emotion analysis. The latter is useful for identifying positive or negative emotions felt by the candidate when answering non-technical questions. Researches show that emotion detection is also possible by using text analysis, based on "emotion words" and being capable of recognizing the eight basic emotions: joy, surprise, expectation, love, anxiety, pain, anger and hatred [28].

In order to ensure relevant feedback, the 3D model representing the avatar must have small changes in its facial expressions, varying from joy to anger, based on the answers received. In the end, a performance module will allow the user to visualize their final evaluation: how many questions were answered correctly, how the soft skills were represented.

The technologies used for creating the chatbot will be integrated in Unity3D. It is a well-known and used game engine, which focuses on portability, being capable of working on a great variety of operating systems. In addition, the applications created in Unity can be exported on different platforms, such as mobile, web, desktop or consoles, allowing also the integration of virtual reality additional devices. Unity uses scripts written in C# or JavaScript and allow the import of 3D models created in well-known 3D modeling programs, such as Blender, Maya or 3ds Max [29].

Regarding chatterbots APIs, there are several available options which can be integrated in Unity and have already been used for creating successful games, such as *Event 0* or *Talos Principle*. SpiritAI, PullString or Alicebot AIML are some example of chatbot AIs systems, with the latter being the most stable and accessible. Artificial Intelligence Markup Language (AIML) is a language similar to XML, easy to learn, which can be easily used for creating and personalizing a dialogue with a chatbot [30].

The VR-JOB system uses an original solution, with a modular architecture, personalized to better fit the purpose of training software engineers. The knowledge base can be easily extended or modified in order to provide training for any engineering field.



**Figure 4: VR-Job System Architecture Diagram.**

In Figure 4, the system's architecture diagram presents the modules contained in our application. Therefore, the VR Presentation layer contains the UI elements necessary for the user-system interaction (main menu, available domains, input and output fields etc.), as well as the 3D Avatar, which reacts based on the user's responses. The Chatbot API layer contains the Alicebot services integrated in Unity, which are trained for hard and soft skills using a knowledge base. The latter is included in the data layer and contains various programming or social-based domains. It can be easily modified and accessed by the administrator via the data access layer and using the AIML. Free alternatives, such as Gaitobot AIML editor are available and easy to use for writing AIML files [31]. Another important part of the system is the performance evaluation module, which collects the user's responses and saves them in a database. Afterwards, a semantic evaluation module interprets the data collected in the database and gives the feedback to the user, who can evaluate their performance.

Our proposed solution does not exclude the integration of other computer-based technologies, such as augmented reality or videos of successful or failed interviews. After the finalization of the initial prototype, quantitative and qualitative indicators will be used in order to evaluate the success or the failure of the system, using as target group software engineering students, teachers and employees.

**CONCLUSION**

In conclusion, the evolution of artificial intelligence and virtual reality can facilitate the development of revolutionary applications for training. Preparing for an interview can cause a lot of stress for many people, especially engineers. By creating a realistic virtual



environment and using an “intelligent” chatbot as a conversational partner, candidates can practice and improve their knowledge and competencies before the actual interview. However, chatbots shouldn’t be viewed as tools meant to replace traditional interviews and the human factor, they are simply a method of training in order to improve one’s performance. In fact, training represents the key to success and virtual reality, the opportunity of achieving this success.

## REFERENCES

1. Brown, M., Engineering Job Market Growth is Slow Entering 2016, (2016), <http://www.engineering.com/>.
2. Harris, S., How to prepare for your first engineering job interview, (2015), [www.theengineer.co.uk](http://www.theengineer.co.uk)
3. Business Dictionary, Definition of word “training”, [www.businessdictionary.com](http://www.businessdictionary.com)
4. Moldoveanu, A., Introduction to VR. *Lectures at Faculty of Automatic Control and Computer Science, University POLITEHNICA of Bucharest* (2016).
5. Heeter, C., Being there: the subjective experience of presence. *Presence: Teleoperators & Virtual Environments* (1992), 262.
6. Pan, Z., Cheok, A. D., Yang, H., Zhu, J., Virtual reality and mixed reality for virtual learning environments, *Elsevier Computers & Graphics*, 30 (2006)
7. Pensieri, C., Pennacchini, M., Overview: Virtual Reality in medicine, *Journal of Virtual Worlds Research*, (2014)
8. Alluisi, E., The Development of Technology for Collective Training: Simnet, a Case History, *Human Factors*, 33, (1991), 343-362
9. Rahimian, F. P., Arciszewski, T., Goulding, J. S., Successful education for AEC professionals: case study of applying immersive game-like vr interfaces, *SpringerOpen Visualization in Engineering* (2014).
10. Chan, J.C.P., Leung, H., Tang, J.K.T., Komura, T., A Virtual Reality Dance Training System Using Motion Capture Technology, *IEEE Transactions On Learning Technologies*, 4 (2011)
11. Microsoft Cognitive Services, (2016). <https://gallery.cortanaintelligence.com/Collection/Cognitive-Services-1>
12. Microsoft Azure, Cognitive Services, <https://azure.microsoft.com/en-us/services/>
13. Mauldin, M., ChatterBots, TinyMuds, and the Turing Test: Entering the Loebner Prize Competition, *Proc. Artificial Intelligence 1994*, AAAI Press (1994).
14. Colby, K. M., Hilf, F. D., Weber, S., Kraemer, H. Turing-like indistinguishability tests for the validation of a computer simulation of paranoid processes, *Artificial Intelligence 1972*.
15. Cohen, K. B., Biomedical Natural Language Processing and Text Mining, *Methods in Biomedical Informatics*, Academic Press, Oxford (2014), 141-177.
16. What is the best API to create a chatbot in, <https://www.quora.com/>.
17. Khanna, 7 Ways VR and Training Bots Can Transform Sales Training (2017) <http://www.trainingindustry.com>
18. Kowalski, S., Pavlovska, K., Goldstein, M., Two Case Studies in Using Chatbots for Security Training, Springer (2009)
19. Tricore Human Capital Management, The pros and cons of chatbots in HR management, <https://tricorehcm.com/>
20. Kwei, K. Y., Chatbots in HR: Improving the Employee Experience, *Advanced Work Based Report*, Singapore. (2017)
21. Job Interview Training with Molly Porter, <https://www.jobinterviewtraining.net/>
22. Humm, L. B., Olsen, D., Bell, M., Fleming, M., Smith, M., Simulated Job Interview Improves Skills for Adults with Serious Mental Illnesses, *Stud Health Technol Inform* (2014), 50–54.
23. Smith, M. J., Fleming, M. F., Wright, M. A., Roberts, A. G., Humm, L. B., Olsen, D., Bell, M. D., Virtual reality job interview training and 6-month employment outcomes for individuals with schizophrenia seeking employment, *Schizophr Res* (2015), 86-91.
24. Klinger, E., Bouchard, S., Légeron, P., Roy, S., Lauer, F., Chemin, I., Nugues, P., Virtual Reality Therapy Versus Cognitive Behavior Therapy for Social Phobia: A Preliminary Controlled Study, *CyberPsychology & Behavior* (2005), 76-88.
25. Kwona, J. H., Powellb, J., Chalmers, A., How level of realism influences anxiety in virtual reality environments for a job interview (2013), 978–987.
26. Hone your Interview Skills with the Virtual Job Interview Simulator, <https://www.eonreality.com>
27. Brank, J., Grobelnik, M., Mladenić, D., A Survey Of Ontology Evaluation Techniques, *Department of Knowledge Technologies Jozef Stefan Institute, Slovenia*, (2005)
28. Quan, C., Ren, F., Sentence Emotion Analysis and Recognition Based on Emotion Words Using Ren-CECps, *International Journal of Advanced Intelligence*, 2 (2010), 105-117.
29. Jasani, T., The top 10 engines that can help you make your game, *The App Guruz Ceo/Founder* (2014).
30. Peake, A., (2017) Chatbot AIs in Unity with AIML and SpiritAI, <https://gamebridgeu.wordpress.com>
31. GAITObot, aiml chatbot hosting, <https://www.gaitobot.de>



# Tower Defense with Augmented Reality

**Bogdan-Mihai Păduraru**  
Faculty of Computer Science,  
“Alexandru Ion Cuza” University of Iasi  
General Berthelot, No. 16  
bogdan.paduraru@info.uaic.ro

**Adrian Iftene**  
Faculty of Computer Science,  
“Alexandru Ion Cuza” University of Iasi  
General Berthelot, No. 16  
adiftene@info.uaic.ro

## ABSTRACT

The current project presents the creation of a game where the interaction is based on standard elements like keyboard and mouse and also with cardboard drawings used to represent the game map. The main aims of this project were both to offer a better user experience for the game players and to help students or interested persons to understand better how the A\* algorithm works. The idea of this project came when Augmented Reality applications started to appear on the mobile market, offering a different experience than current standards.

## Author Keywords

Augmented reality; Unity; A\* algorithm.

## ACM Classification Keywords

H.5.2. Information interfaces and presentation (e.g., HCI): User Interfaces. H.3.2. Information Storage and Retrieval: Information Storage.

## General Terms

Human Factors; Design.

## INTRODUCTION

We live in an era where stress affects us more and more every day. Relaxation by playing video games has become a normal activity for most of us. Most people now own a smartphone or tablet which they carry it everywhere [3]. These devices contain applications that help them stay in touch with friends or family (Facebook, Skype, chat apps), visualize multimedia content (movies, music, online TV) and also video games that must be up to date with the consumers need of relaxation and spending small breaks time.

This project will present an application that was built with the help of Augmented Reality with its purpose being to offer a relaxation alternative for those who use it. Right now the number of similar applications is pretty small, the most popular one being AR Defender2.

In what follows, the paper is structured as follows: a state-of-the-art section where similar application that use Augmented Reality are presented, a section for system architecture where it shown the basic elements of the system we build, and a final part where it shows conclusions and suggestions for future work.

## STATE-OF-THE-ART

The world of Augmented Reality started to be more and more present in applications for tablets or smartphones. We will now present the most important applications that are similar to our application.

### AR Defender 2

This application was one of the first to come on the mobile market, being developed and published by Bulkypix at the beginning of 2013 [1]. By mixing augmented reality components with 3D objects that represent different structures and characters inside the game this application immediately drew the attention of users. One of the features that made this application so successful was the networking that made it possible to play together with friends at the same time.

### PulzAR

Initially published as a single game by the existing developer Exient Ltd, this application was later converted into a package of games, all of them using Augmented Reality technology [5]. One key feature this application had when it launched was the platform where the games could be played, PS Vita. Another interesting thing to notice was the launch date, 12 June 2012, only four months after the initial launch of the PS Vita console.

### Toyota 86 AR

Another application that uses Augmented Reality technology, Toyota 86 AR was developed by Toyota Company in order to showcase the driving experience of their newest car model [6]. Made for Android and IOS, this application contains a single scenario where as soon as one of the games drawing maps made by Toyota are recognized, the user can start controlling the Toyota 86 car model in a virtual environment which serves as a real simulator for those who want to inspect and test the form, components, acceleration ratio, drifts etc.

## SYSTEM ARCHITECTURE

In this section the main objective is to give a better understanding of the events that occurs in the current application, to familiarize with some of the used concepts and to offer some details about their implementation.

## Game matrix

The first thing that must be prepared once the level starts is to build the game map both visually on the screen and internally in memory. Because in each scene there are only two objects that must be recognized by the Vuforia SDK [8], we can hold direct references to them. Once we know the object on which the map will be built, we can access its components and get details about the map size with the `GetSize` method that return a bi-dimensional vector where we find the width and height. Another important thing that we must be aware of is the dimension of a 1x1 piece of the matrix in relation with the world space. For this we have created an enum structure for an easier configuration, leaving the user to make the decision before the game starts.

Once this is done, we must decide where to place the start and finish point (represented in the image with “S” and “F”) on the XY axis in our map from a visually point of view. Let’s consider the rectangular shaped below as one of our game drawings where we will build our map. One more thing to remember is that each object in Unity is represented by a position vector that describes the coordinates of its center [7].

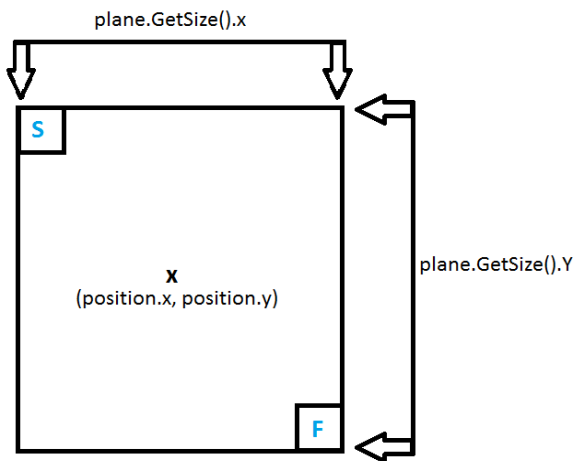


Figure 1. General representation of a game map

The only thing that remains is to create the matrix that will correspond to the game map. Since the A\* script contains the functionality of finding an optimal path based on a given map, it will be responsible with the matrix creation and all its details.

### Detecting the map drawing

For the map drawing detection in a level we have used the `OnTrackingFound` method that is found inside the `CustomEventHandler` class which was made available by the Vuforia SDK. This script is attached to an object that

will soon be detected and offers some details about the object, like its name by using an instance of the class `TrackableBehaviour`, called `mTrackableBehaviour`.

Sending the message when an augmentable object was found is done like this:

```
MapCreator.instance.MapFound(mTrackableBehaviour.TrackableName);
```

Once this message was sent, we continue with adding obstacles for our current level and then, we search for the optimal path for our enemies that will soon be instantiated. We can see in Figure 2 how this map looks in our application.

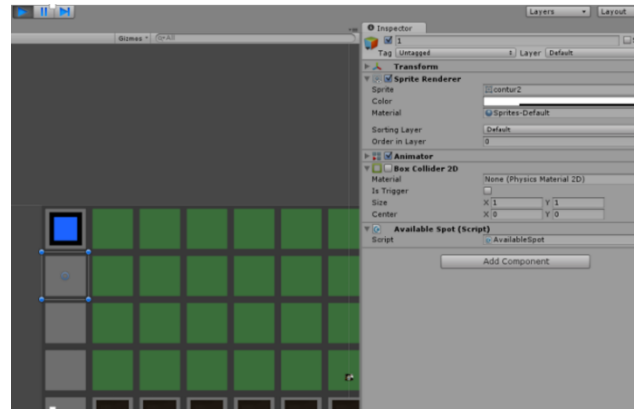


Figure 2. Map space object.

### Adding obstacles for current level

When sending the message that our augmentable object was found, we establish our current level index, after which we start instantiating obstacles on the map.

Using two “for” instructions from 0 to the maximum number of lines respectively columns, we will move through each space of our game map. One important thing while doing the obstacles initialization is to register them on our map:

```
Astar.instance.RegisterObstacle(i, j);
```

- where  $i$  represents the index of our current line and  $j$  the index of our current column. In order to instantiate an object we used the `Instantiate` method from the `MonoBehaviour` class:

```
Vector2 pos = GetWorldCoordsFromMatrix  
Coords(new Vector2(i, j));  
  
GameObject c = Instantiate(obstacle, new  
Vector3(pos.x, pos.y, 0), Quaternion.  
identity) as GameObject;
```

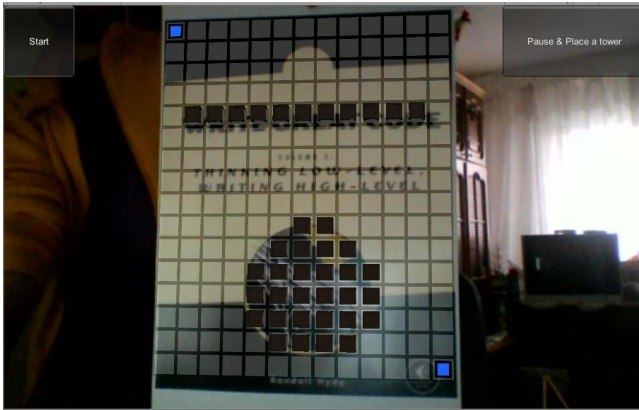


Figure 3. Showing the complete map of a level.

The map from each level is built according to the black portions from the image (see Figure 3). In the matrix representation of the level we say that a space is available when its value is 0, and an occupied space has its value equal to 1. Thus when the pathfinder algorithm will run this representation will help it search for the optimal path and avoid all obstacles.

#### Finding the optimal path using A\* algorithm

A\* is the most popular choice when it comes to pathfinding since its very flexible and can be used in a big range of contexts, which is the primary reason we used it.

The base idea of A\* [4] comes from Dijkstra algorithm [2] because it can find the shortest path, but it also contains elements from Greedy algorithm since it is based on a minimum constant function for guidance during the search, called also a heuristic function. The success of this algorithm consists in combining these two information's from Dijkstra algorithm [2] (which favors nodes that are closer to the *start* node) and Greedy algorithm (that favors the nodes that are closer to the *end* node). In the standard terminology when we discuss this algorithm,  $g(n)$  represents the function that computes the exact cost from *start* to any node, and  $h(n)$  represents the heuristic function that estimates the cost from a current node to the *end* node. The logic behind this algorithm is to balance these two function when moving towards the final node. Thereby, at each iteration this algorithm will chose the solution that has the smallest value which is associated to the function  $f(n) = g(n) + h(n)$ .

In a map represented as a matrix there are numerous heuristic functions for movement, however we will mention only two of them, the first one being used in this project.

First the Manhattan distance that allows movement in four direction in a matrix represented map: *north*, *south*, *east*,

and *west*. This choice is the most popular because it is easy to compute with the following function:

```
function heuristic(node, goal){
    dx = abs(node.x - goal.x);
    dy = abs(node.y - goal.y);
    return D * (dx + dy);
}
```

An important thing to notice is the value  $D$  which is a constant and must be assigned at the beginning. For a shortest path computation and keeping the heuristic function “admissible” it is recommended to assign a small value for this. Besides, there are different proposals related to this constant that take into consideration more aspects of a map like different terrain that imply different traversal costs or the number of the obstacles in a scene, but mostly this constant variable is assigned with 1, representing the simplest case scenario which was also used in this project.

When we need a more detailed traversal of the map we can use eight directions which implies a heuristically function that takes into account diagonal computation. This function has the same effect as the one mentioned earlier when it receives the four classical directions, but it can also support new directions: north-east, north-west, south-east, and south-west. The function that computes the distance between two given nodes with this type of movement is:

```
function heuristic(node){
    dx = abs(node.x - goal.x);
    dy = abs(node.y - goal.y);
    return D * (dx + dy) + (D2 - 2 * D) *
min(dx, dy);
}
```

The constant variable  $D$  was discussed above, but we can see a new variable,  $D_2$ . The value assigned to this is strictly connected to the value of  $D$  with the given formula:  $D_2 = \text{sqrt}(2) * D$ .

Once these details have been discussed we can move on to present the A\* algorithm. There are two data sets, OPEN and CLOSED. In the first set we add nodes that are possible candidates when computing the path. Initially this set contains only the start node. The CLOSED set contains nodes that have already been examined. Initially this set is empty. If we had to give a visual representation of these sets we could say that the first represents the frontier and the second shows the inside area of the visited zones. Also each node will have a reference to its parent or better said to the node through we reached it in

our search, being able to determine the path by going recursively through each parent of a node once we found the final node.

The algorithm consists of a while loop that will pop out at each step the best node from the OPEN set (the node with the lowest value of its f function). If that node is the final one, the algorithm will stop. Otherwise, the node will be deleted from the OPEN set and added to CLOSED. Then we examine all its neighbors. In any of these neighbors is in the CLOSED set, they will be skipped. Also, if one of them is in the OPEN set we skip them, avoiding a double check. If none of these conditions are met, the node will be added to the OPEN set and we will set its parent reference with the current node and compute its cost equal to  $f(n) = g(n) + h(n, n')$ , where  $n'$  represents the extracted node.

```

OPEN = priority queue containing START
CLOSED = empty set
while lowest rank in OPEN is not the GOAL:
    current = remove lowest rank item from
    OPEN
    add current to CLOSED
    for neighbors of current:
        cost = g(current) +
movementcost(current, neighbor)
        if neighbor in OPEN and cost less than
g(neighbor):
            remove neighbor from OPEN, because
new path is better
            if neighbor in CLOSED and cost less
than g(neighbor): **
                remove neighbor from CLOSED
            if neighbor not in OPEN and neighbor
not in CLOSED:
                set g(neighbor) to cost
                add neighbor to OPEN
                set priority queue rank to
g(neighbor) + h(neighbor)
                set neighbor's parent to current
reconstruct reverse path from goal to start
by following parent pointers

```

In this project the OPEN and CLOSED set have been represented as simple generic lists that contains element of type *AstarNode*, which is a data structure create by us with the following representation:

```

public int x;
public int y;
public AstarNode parrent;
public double g;
public double rank;

```

where “x” and “y” represent the coordinates from the game matrix, “rank” represents the value of the function  $f(n) = g(n) + h(n)$ , the rest of them being easily to figure out from their name.

In order to help the players of the game, all details of this algorithm can be presented to the students in class before using this application. In the future we intend to add an option to the application, where the players can see these details of the A\* algorithm.

### Starting the game and details about the game mechanics

Once the first six activities from the beginning of this chapter have been successful, in the top left corner of the screen a button named “Start” will appear, its role being to start the game (See Figure 3).

Since the last activities refer to the start and end of the game, something that is knows through the *EnemyController* class, this one will be responsible the send messages to the *UIManager* class to update the screen interface according to the game state. Once the game has been started the *EnemyController* class will handle the instantiation and management of the enemies. When an enemy has been created he will receive a confirmation message that will initialize all of its data, and after that it will be registered in a list of active objects of the class that created it. After all this setup has been done for an enemy, the class will launch the enemy activity. Each enemy contains two main components, *UnitMovement* and *EnemyStats*.

The first component will send a request to the A\* script for the optimal path from the starting point to end point, and it will received a list of Vector2 objects that contains the coordinates that must be followed step by step. Once this list is obtained the object will start moving from its initial position to the next point from the list. For the object movement we have used a function which is found in the *MonoBehaviour* class, *Vector3.MoveTowards()* which has the following signature:

```

public static Vector3 MoveTowards(Vector3
current, Vector3 target, float
maxDistanceDelta);

```

The first parameter represents the initial position from which we start, the second represents the final position that we want to reach, and the third parameter represents a measuring unit that will point the object towards to destination, or for a better understanding it can be viewed as a speed that dictates how fast the object reaches its destination. In order to move an object with this function

we must use it inside the *Update* function (function that is being called at the beginning of each game frame).

The second component, *EnemyStats*, contains details about the object life time. When an enemy is attacked it will lose points from its current health and once this value reaches below zero he will autodestruct, but not before sending a message to the script that is responsible for the enemy management.

Once an enemy started its activity the only interaction he will have will be with the tower components and their projectiles represented through the *TowerLogic* and *BulletLogic* classes. These contain functionality for physical interaction between object with the help of the *OnTriggerEnter* and *OnTriggerExit* functions.

Another activity, that takes place after the game is started, is the tower placement in one of the free spaces of the map (see Figure 4). For this we have used another function from *MonoBehaviour* class, *OnClick()*. This function gets called each time the user physically presses the click button of the mouse and the cursor is over the object that contains a script that has this function and a collision component, which in our case is a *BoxCollider*. For this we have built a new class called *AvailableSpot* that is attached to each object that represents a piece of our game map. Once the obstacles have been added and the enemy path is known, the collision component will be disabled for the objects that meet these conditions so that the player can no longer interact with them since they are occupied spaces. As we can see from the picture below that the green spaces are available for building towers and by selecting one of the occupied spaces we find the collision component (*BoxCollider*) inactive.

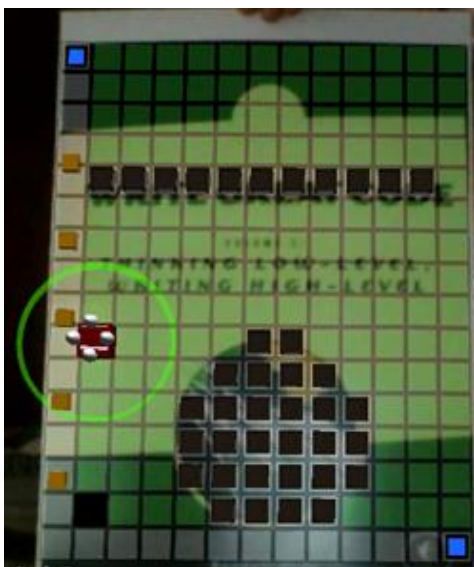


Figure 4. Displaying available spots and a tower

At this step, we can use different possible configurations for the added towers, which can help the students to understand better how the algorithm A\* works. There are cases when after adding a tower it is possible not to find a path from the *start* point to the *end* point, there are cases when the added tower doesn't influence the initial solution of the A\* algorithm, or there are cases when A\* algorithm build another solution. All these cases must be analyzed and discussed with the students.

### Description of a level and necessary steps to finish

In this section we will present one of the built levels and all the steps that are necessary in order to finish it.

The top-left and bottom-right corners represent the *start* and *end* point of the level. By pressing the *Start* button the enemies will be instantiated, and their main purpose is to reach the end point. Our goal is to prevent them from doing this. By activating the "Pause & Place a tower" the game will be paused and the user can select one of the available spaces, represented by green, where he can place towers that will destroy the enemies. Once a spot has been selected, a confirmation button will appear bellow, after which the game can be resumed or return to the state where it waits for the user to repeat the action we just mentioned.

Once all enemies have been eliminated, a message will be displayed on the screen and we can return to the main menu scene by pressing the "Escape" key (See Figure 5).



Figure 5. Finishing a level

### CONCLUSIONS

As mentioned in introduction, the aim of this project was to create an educative game including elements of Augmented Reality. Matter of our using Unity, the Vuforia SDK's and developed algorithms we managed to



develop an app that offer the possibility of transforming a part of it in real world game space. Thus users can increase interest since the board on real interaction represents a degree of novelty for attracting Player. Also, the user can add obstacles on the map and create different possible configurations, and in this way he will understand better how the A\* algorithm works.

This application can be extended to other platforms, but also we can add new features such as connecting multiple participants to the same game session, creating new game mechanics such as upgrading towers, creating multiple terrain types, each with a cost different to be covered, adding new models for representing objects in scene and even the ability to add their own drawings with easy to use editing interface obstacles.

When we first started this project, we did not have in plan a different interaction for the players, but as we continued working we soon realized that this type of interaction would give the game a better feel and experience.

#### **ACKNOWLEDGMENT**

This work is partially supported by POC-A1-A1.2.3-G-2015 program, as part of the PrivateSky project (P\_40\_371/13/01.09.2016).

#### **REFERENCES**

1. AR Defence 2: <https://itunes.apple.com/en/app/ar-defender-2/id559729773?mt=8> (Last time accessed on 29 May 2017).
2. Dial, R. B. Algorithm 360: Shortest-path forest with topological ordering. *Communications of the ACM*. 12 (11), (1969), 632–633.
3. Digital Marketing Megatrends 2017: <http://www.smartinsights.com/mobile-marketing/mobile-marketing-analytics/mobile-marketing-statistics/> (Last time accessed on 30 June 2017).
4. Hart, P. E., Nilsson, N. J., Raphael, B. A Formal Basis for the Heuristic Determination of Minimum Cost Paths. *IEEE Transactions on Systems Science and Cybernetics SSC4*, 4 (2), (1968), 100–107.
5. PulzAR: <https://goo.gl/3FWM8n>. (Last time accessed on 29 May 2017).
6. Toyota 86 AR: <http://www.toyota86ar.com/>. (Last time accessed on 29 May 2017).
7. Unity: <https://unity3d.com/> (Last time accessed on 29 May 2017).
8. Vuforia: <https://developer.vuforia.com/> (Last time accessed on 29 May 2017).

# Gamification in the context of user productivity and engagement

**Sfât Lorena Adina**

Computer Science Department  
Technical University of Cluj-Napoca  
Cluj-Napoca, Romania  
sfatlorraine.03@gmail.com

**Teodor Ștefănuț**

Computer Science Department  
Technical University of Cluj-Napoca  
Cluj-Napoca, Romania  
teodor.stefanut@cs.utcluj.ro

## ABSTRACT

This article presents an approach on dealing with the engagement of mobile devices users, with the purpose of increasing productivity. As it seems that more and more smartphone users acknowledge the benefit of tracking down things they do, from physical to social activities and doing so in a competitive way, a lack of solutions that address a wider range of coverage was noticed. The proposed solution provides motivation and a competitive environment for users, while tracking their objectives and challenging others, by using an approach that covers the best practices of the concept in use and by being developed in a user-centered manner. Unlike existing applications, our proposed model covers a wider range of objectives and encourages the engagement in a game-like manner. The application's fundamental background consists of the concept of gamification, applied in a day-to-day environment.

This concept, together with a series of other user-oriented concepts implemented in the proposed application, have as main goal maintaining users' interest while ensuring increased productivity. This application is meant to cover a wide range of necessities, for smartphone users, by allowing users to track objectives proposed by themselves or others, as well as ensuring motivation through a gamification system.

## General Terms

Gamification, motivation, engagement, SMART, usability, user-centered design

## INTRODUCTION

Our everyday life is becoming more and more loaded, and we end up with a full agenda. The application proposed in this paper promotes an easier way to track one's objectives, with the use of a device that is almost never left out of sight, the mobile phone. By doing so, we attempt to create a solution that maintains the users' interest, increases their productivity and always challenges them to exceed others or their past achievements.

In this paper, we propose an application that motivates users to accept challenges and complete objectives in

a competitive way by providing a user-friendly interface.

While there are many applications that serve as agendas, their lifetime is significantly short. It does not take much for the user to slowly forget about them or lose interest and ending up deleting them soon afterwards [1]. Therefore, we propose a user-centered design and as a main tool, the gamification technique.

In large terms, gamification refers to the application of game-design elements and game principles in non-game contexts. Such elements and principles are adopted in order to increase motivation, improve user engagement, ease of use and even usefulness of systems. The idea came from the fact that people like having clear goals and even more being awarded once achieved. In this way, the users' behavior is easily influenced to do the expected actions when treated as "players" with game elements such as challenges and rewards [2].

The roots of gamification lie in the User-Centered Design, considering the user in every decision made. User-Centered Design (UCD) means, more than anything, understanding exactly what users need, how they behave in given circumstances and what is their approach - and considering all these in the process of development [3].

As mentioned before, a fundamental step of the gamification process is considering the users. In most of the cases, the users do not have the same complex overview of the project as the developers, thus they will focus more on the fun than on the complexity. In this way, the features should be filtered based on how intuitive and understandable they are for the users.

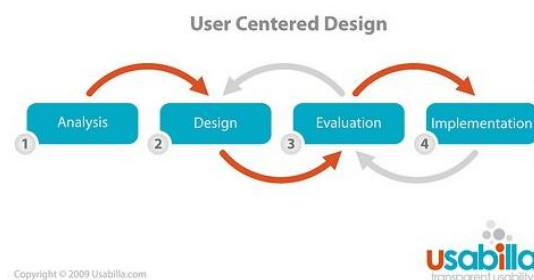


Figure 1. User Centered Design Diagram [3]

When leading on a UCD, one focuses more on

analysis, design, (re)evaluation and implementation, visually illustrated in Figure 1.

Gamification uses its own version of UCD, namely Player-Centered Design (PCD). The user from the standard User Experience (UX) design [4] now becomes the player. It follows the same pattern, by placing the player at the heart of the process. The PCD starts from the premise that the “players” in the exercise are participating because they want to, unlike the “users” from the UCD.

The main purpose of gamification is to engage, motivate and encourage the audience to exceed previous results or other competitors. Players, unlike users, have the free choice of getting involved and it is this option that enables them to have fun.

One of the main concepts that enhance the engagement of users are the challenges. Human nature makes us accept challenges and prove our ability to handle them. In order to enhance its effect, it is a good idea to use rewards in order to increase the motivation as well.

Competition is also very stimulating. Leaderboards increase significantly the desire to be the leader, and by stimulating the enthusiasm, it engages the user.

By making a promise that awards will come to those who accomplish tasks, the users are encouraged to go back and the curiosity, excitement and competitiveness drive people to spend more time on the app. The idea that completion of an activity has as result getting something in return keeps you motivated and engaged. It is all about tricking the mind, and at a point when it can get boring, the system should be able to engage again, by playing with psychology.

Setting goals is one of the fundamental aspects of a good gamification system. You can see in the Figure 2 the meaning of the SMART acronym [5].



Figure 2. SMART objectives [6]

In order to make a goal specific, one can simply answer to the “Who, What, Why, Where” type of questions. By making it measurable, you make sure you are heading towards completing a goal. Attainable means being realistic when setting a goal and making sure it is achievable. Relevant is relative to you, making sure you like it and formulating it in a way that it challenges you. One last thing to set is the time period and making sure the goal is time-bound and so taking advantage of the effect time has on decision making. Setting SMART goals works like a contract between parties, by ensuring they use the same

terminology, that they agree upon the measures.

However, it was proven countless times that the reality is never as perfect as a plan on paper. While it can be easy to think of the benefits of establishing SMART goals, in real life it can be hard to be specific, measurable and time-bound [7]. It is easy for users to get stuck in front of the input boxes and not know how to adapt their specific situation. The issue is even more visible in the moment when the objective is not fully understood. Sometimes one cannot assess how much time it would take to achieve goal X simply because it is too far in the future or has no experience in the field. It is these exceptional cases that might come in the way of users achieving full engagement.

Leaderboards that were mentioned before, work hand in hand with goal setting. By achieving certain goals, one can advance in the leaderboard and receive afferent rewards.

The concepts above have been integrated in the application presented in this article, which is currently available, as a proof of concept, on Android devices. Platform selection has been related mostly with subjective authors preference, and has very little to do with specific capabilities required for the application domain itself.

The ideal scenario would require less input from users and would use existing APIs to keep track of certain categories of objectives.

This paper focuses on presenting the concept dealt with, such as game mechanisms, game principles and engagement methods, presents related work in direct comparison with the proposed application, a psychological side of motivation followed by a brief presentation of the specific implementation of the concepts in use.

## RELATED WORK

Google provides an application, similar in purpose with our proposed application. However, unlike our solution, Google Fit [8] targets a more specific area of motivation. It is oriented towards sports or fitness motivation, helping users keeping track of their objectives (e.g. running, walking). The application has a small number of predefined objectives, but these are too restrictive, focusing only on the ones mentioned before. Users are not allowed to see other users’ accomplishments so the only competitiveness present here is the one with past self.

Another similar application is Wunderlist [9] which is ideal for keeping track of things, in the form of a list. The application comes with a series of predefined lists, but they have limited space for information. It is more oriented to tracking shopping lists or lists of movies, for example. These lists can be shared; thus, users can have an overview of other persons’ activities. However, there is no competition, the access to information being strictly informative. One could use this to track short term objectives, that do not require

a lot of follow-up, but it is not suitable for a bigger objective that needs milestones in order to track it properly.

In this way, a need for an application that not only helps users tracking goals, but also engages them by consolidating a competitive environment with the purpose of increasing productivity was noted. This application is suitable for personal use, but can also manage a more professional environment, such as a workplace, where team members can challenge each other.

## **GAMIFICATION CONCEPTS**

### **Game mechanics**

The whole purpose of gamification is to educate a desired behavior of users. Thus, the first step of building such a system is to set an overall goal and the expected behavior [10].

Next, one should consider how to motivate the users to act as expected. While it is easy to motivate users for a short term, finding valuable and invested users can be more demanding. This is where the difference between intrinsic and extrinsic motivation comes in. Extrinsic rewards can be easily demotivating. By giving rewards for the completion of a task only to certain competitors, the others feel unsatisfied and predisposed to lose interest in the service provided, thus the motivation was short-termed and can lead to loss of users. Intrinsic motivation, however, addresses the user directly, by rewards specifically customized for him/her, by referring to more abstract accomplishments, such as efficiency, productivity, personal achievement, sense of pleasure and satisfaction, as well as more specific emotions such as loyalty.

As a last aspect, the motivation should be supported by motivators. These are rewards. Depending on the type of application, it handles one type of economy, be it financial or game-invented. The most typical elements are reward points, badges, leaderboards and other forms of charts.

### **Game principles**

Gamification principles are all about the game design [11]. Applications started using game principles because games are engaging and promote continuous use. Thus, these game principles have been borrowed by applications that have few or no things in common with games, in order to fuel the motivation and attention of users.

One first principle is letting users know how they are doing. Use visual indicators to make their status and their progress visible and to illustrate how they are getting closer to the goal. By seeing progress, users are more motivated to continue.

Another principle is the rewarding system. When users complete a goal, we reward them. It could be either a leveling up, an increase of score, going up in a leaderboard or simply a visual change, depending on

the context.

The look and feel of an application, while not being the most important aspect, it has a significant impact on the users' interest and engagement. Aesthetics is crucial since the expectation of mobile phone users have increased significantly in the last years. In this way, the visuals of an application make up the first impression of the app. The look and feel is what sparks the initial interest and then drive the engagement.

By giving the users something to aim for or something to collect, a great motivation can be achieved. This is the core of user satisfaction since it allows users to relate to others, by sharing and comparing results.

### **Competitive engagement**

Even though people grow up and become functional members of society, they never stop wanting to play. Games have a factor that stimulates and keeps players engaged. The main trick is competitiveness, either with past self or with other players. Almost always, the scope is collecting. Collecting reward points, badges or money in a specific economy.

Games are mainly associated with competitions. Highest performances are rewarded, motivation coming either intrinsically or extrinsically [12] - motivated by doing our best to reach a goal, by our work getting there or by external factors such as badges, collector's points or monetary rewards.

Game factors found in non-game applications have the purpose to enable what we call friendly competitions. By always being able to compare themselves with others, users are more likely to continue using the application and coming back to it in order to exceed past results or others. With the help of the leaderboard, the user has a sense of omniscience and transparency, always knowing exactly who they have to surpass and how much they need to gain in order to do so. By visually seeing how they climb up the ladder, they become actively engaged.

### **Motivation**

We settled so far that there are two categories of motivation for an application's users, namely intrinsic and extrinsic motivation. The end scope of providing these motivations is emphasizing lasting engagement.

Extrinsic motivation means acting towards a goal with the scope of receiving external rewards such as points, badges or money in some monetary system. For users that are extrinsically-motivated, it is not the action or behavior of getting there which they like, but the outcome. Meaning that they might not like what they are doing, but continue doing it for what they get in return [13].

On the other hand, intrinsic motivation focuses on internal rewards like the sense of satisfaction, positive feelings and happiness. Unlike extrinsically-motivated people, these people have a genuine desire to achieve the goal, while enjoying the process itself and having

a great time doing it.

Before considering a user engaged, he/she is no more than an observer. An observer is prone to lose interest. However, if you can reach them, they become engaged, evolving into active “consumers” who will persuade their goals and will head towards our desired behavior. This is precisely why the main focus must be on motivation. Users must be motivated otherwise the audience will become a crowd of disinterested, disengaged and demotivated people, losing sight of the app’s values, highly prone to stop using it.

**PROPOSED APPLICATION**

**Overview**

Before going any further, there are a few technical details to be noted. Two of the main concepts of the application are the challenges and the objectives. Objectives are proposed by users for themselves. They set up a deadline, a description, a suggestive title and other helpful information. Challenges are proposed by users and are available for all the others as well. A challenge, once accepted, turns into an objective for that user. The challenger himself can accept the challenge.

From a technical point of view, the system consists of an API and an Android application serving as Graphic User Interface. A brief presentation can be seen in figure 3.

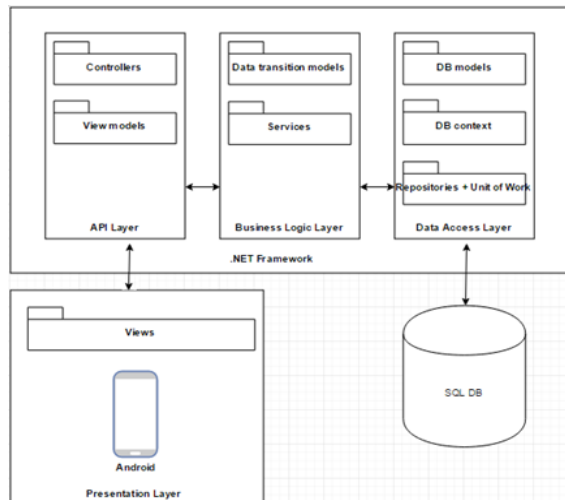


Figure 3. System architecture

To increase usability, users are faced with multiple ways to initiate objectives. They can either launch an objective themselves, or accept a challenge launched by another user. To keep things as structured as possible, challenges follow the pattern of SMART goals to describe what is expected from users as a result, after finalizing the challenge.

In order to emphasize the concept of gamification, users are able to compare results with others and his own old statuses, in a page of reports, with highly visual slots.

The usability is also increased by the notification system implemented using the cross-platform solution provided by Google, Firebase Cloud Messaging, by the user-centered design and by the competitive algorithm maintaining ranks between users. Besides this, the proposed solution is implemented in a manner that makes the user only one tap away of the required information.

**Proposed game mechanics**

By considering the game mechanics presented before, a different approach to the rewards system was adopted, one that tries to provide a better balance between the intrinsic and extrinsic motivation for its users. The image below shows the key elements of the gamification system.

Expected Behavior	Motivation	Motivators
Track objectives and accept challenges	Challenges	Competition with other users; Reminders; Challenger grade
	Loyalty	Custom / Personalized rewards
	Rewards	Leaderboard; Progressive icons; System grades
	Ranking	System grades; Users grades
	Purpose	Increase productivity; Track objectives

Figure 4. Gamification elements

They work as follows: the motivators support the motivation needed by the users to achieve the desired behavior.

This system emphasizes competition between users by encouraging them to accept challenges proposed by the other players. The supporters are then grades received from the system upon completing a challenge/objective, grades received from other users and, optionally, grades received from the challenger. The progress is also represented through a progressive set of icons that are gained upon achieving certain steps. It encourages loyalty by providing custom and personalized rewards upon reaching certain milestones.

The application also comes with a notification algorithm that is implemented in a way that engages the user even more, by providing relevant notifications and avoiding overwhelming users with not-so-relevant messages. For example, users are prompted when someone accepted their challenge for the first time or when an objective is about to reach its deadline.

The end scope of this is the ability to track the progress on user’s objectives, with the purpose of increasing productivity and engagement.

**Proposed game principles**

By now, we know that game principles are basically borrowed from games’ psychology in order to increase engagement and to increase continuous use.

To start with more general aspects, the looks and feels of the proposed application follow the KISS [14] principle, acronym for "Keep it simple, stupid". It uses a flat design with a simple color palette, allowing users



to focus on content while also providing catchy elements that emphasize the game-like aspect of the app. The application is consistent through its screens, maintaining the looks throughout the whole application. Sample pictures from the app can be seen in Figure 5.

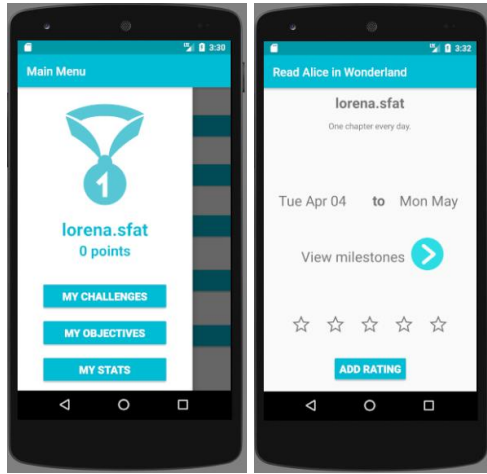


Figure 5. Application User-Interface sample

The application also proposes a ranking algorithm that works hand in hand with a rewarding system. Users gain their rank as an arithmetical mean between the grades of all their finalized objectives. We call objectives finalized when they were either completed or are past deadline. When an objective is finalized, through any of the means mentioned before, the system automatically gives it a grade, based on the difference between the deadline and the completion date or the lack of it. Penalties are applied and are calculated considering the duration of the objective, the start date, the end date and how long the deadline was violated. For example, if I planned to learn 3 new words in a foreign language, every day, for 2 weeks, but I started one week later and finished one week after deadline, I still held up part of the bargain, so the penalty will be smaller than in the case I simply prolonged the implementation with one week. In case of an accepted challenge, that is a challenge turned into an objective, the system also considers the correlation between the deadlines chosen for planning steps by the challenger and the actual end dates. Penalties for surpassing deadlines for milestones are also applied, by the same principles as for the objectives.

With the use of graphical, visual elements, just like in a game, users can easily track the status of their challenges and objectives as well as their progress and their rank. Users can filter through the list of objectives to see the ones that are completed, overdue or to keep track of how many other users accepted their challenge, all these strategically and intuitively available for the user. Users are also provided with a screen for statistics, where they can track their progress compared to other users and they can also see statistics, personalized for their accomplishments - for example the most popular challenge, the best-graded

objective.

The proposed application indicates to users their progress visually.

### Proposed engagement methods

The main engagement method that we propose is the “competitive gameplay” [15]. We challenge the users to score as high as possible, to surpass their competitors and to do their best while finishing a challenge in order to get the best grades.

We engage the users by making it easy for them to have an overview of their challenge and/or objective stats through a specific screen with leaderboard and statistics. To help them keep track, users are visually notified when objectives are about to reach the deadline, by receiving a general message on the main screen. When entering the list of their objectives, the priority is set to the ones that are about to reach the deadline sooner.

To keep them focused on the challenges they launch, users also have access to: information about users accepting those challenges (e.g. number of involved users), a general status overview when players finish a challenge implementation and the possibility to add a review for a player. Remember that this review is a part of that objective’s grade, together with the grade from the system.

We know from experience that users are prone to forget about an application, simply because of a busy schedule, leaving aside cases where the app does not meet their needs. We need to remind users about the app, and the best way to do so is implementing a notification system. To emphasize what we said before about this algorithm, we point out the main types of notification:

- with the purpose of reminder, the user is notified when he/she hasn’t used the app in a while.
- with the purpose of engagement, as updates, we notify the user when they reach a significant position in the leaderboard, when a challenge they launched becomes popular etc.

The whole structure of the application focuses on the user’s needs. Everything is placed in a way that the user is always one tap away from getting to the desired screen or performing the desired action.

For example, we considered intuitively to place the more accessed screens on a “drawer menu”. This menu is always available by swiping on the screen. There, users can access their statistics or objectives pages or access the list of challenges they launched.

### APPLICATION EVALUATION

There are multiple ways of evaluating a system. Considering our proposed application, the evaluation will assess aspects on usability and also on performance improvements of the involved users in the contexts of the accepted challenges. Therefore, we propose a user-based evaluation technique.

Data acquired through the evaluation will give an insight on which further improvements that will enhance the ease of use and, specifically, increase user's performance and number of achievements, are recommended.

Designed experiments address two different but complementary environments: work and leisure. The application can either be integrated in the work process of a team or used between a group of friends, to track and compare common objectives/challenges. Conducted measurements will provide detailed statistics about the number of achieved objectives in a set interval of time (e.g. a week, a month, a year). In the same way, we will be able to track the average time a player needs to achieve an objective.

After collecting this data, we must interpret the results. We will compare diverse groups of users based on certain criteria (e.g. those who used the application for X months vs. those who haven't used the application almost at all) to see the factors that motivate or discourage players. The data will be interpreted for singular user and groups of users based on past results.

The types of users that take part in these studies have a significant impact on the results. For example, a user that never uses the mobile phone except for making calls, will be less likely to be engaged by any kind of application. In the same way, the reaction also depends on whether users are familiar with the concept of gamification, their types of objectives and challenges, whether they are "team players", etc.

## CONCLUSIONS

This approach intends to cover best practices in the domain of gamification in the context of an application that focuses on users' productivity and has as goal users' engagement.

The proposed application is built in a user-oriented manner, having all screens and activities strategically and intuitively placed at hand. The only thing required from the user is the ability to create objectives and a challenge in a SMART way and being objective and fair to other users to maintain a satisfying environment.

## Future development

The next step of our research will be focused on testing the performance of the application in different environments and provide support for more versions of this operating system, while also fixing issues that were not dealt with so far.

Considering the fact that this application focuses on increasing users' productivity, a feature that would be next on the list is one that measures this possible improvement. Be it either by simply asking for

feedback or through a more complex system. Together with the results from the evaluation, such a system could help building up statistics that would help the developers know where to invest more time, what to change to please the user and to have an overview from a non-technical point of view.

## REFERENCES

1. Schlägerová, P. (2016). Top reasons why users lose interest for your app. Retrieved from Leadmill: <http://www.leadmill.com/top-reasons-why-users-lose-interest-for-your-app/>
2. Kumar, J. M. (2016). Get to Know Your Players for Your Gamification Project. Retrieved from Interaction Design Foundation: <https://www.interaction-design.org/literature/article/get-to-know-your-players-for-your-gamification-project>
3. Kumar, J. M. (2014). User-centred Design in a Gamification Context. Retrieved from Interaction Design Foundation: [www.interaction-design.org/literature/article/user-centred-design-in-a-gamification-context](http://www.interaction-design.org/literature/article/user-centred-design-in-a-gamification-context)
4. Hoßfeld, T., Heegaard, P.E., Varela, M. (2016) - Quality and User Experience. doi:10.1007/s41233-016-0002-1
5. Doran, G. T. (1981). "There's a S.M.A.R.T. way to write management's goals and objectives". *Management Review*. AMA FORUM. 70 (11): 35–36
6. Turkay, S. (2014). *Setting Goals: Who, Why, How?* Harvard Publications
7. Goerg, Sebastian J. (2015). "Goal setting and worker motivation". *IZ A World of Labor*
8. Google Fit, <https://www.google.com/fit/>
9. Wunderlist, <https://www.wunderlist.com/>
10. Andrzej Marczewski (2013). *Game Mechanics in Gamification* <https://www.gamified.uk/2013/01/14/game-mechanics-in-gamification>
11. <http://datagame.io/gamification-principles/>
12. Ganit Richter, Daphne R. Raban, Sheizaf Rafaeli (2015) - Studying Gamification: The Effect of Rewards and Incentives on Motivation
13. Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does Gamification Work? – A Literature Review of Empirical Studies on Gamification. In proceedings of the 47th Hawaii International Conference on System Sciences, Hawaii, USA
14. Eric Partridge, Tom Dalzell, Terry Victor (2007) - *The Concise New Partridge Dictionary of Slang*, Psychology Press, p.384.
15. Hee Jung Park, Jae Hwan Bae (2014) - Study and Research of Gamification Design

# Conversational Agents Embodying a Character Using Neural Networks

Mihai Ilie and Traian Rebedea

University Politehnica of Bucharest, Faculty of Automatic Control and Computers

Bucharest, Romania

ilie.mihai92@gmail.com, traian.rebedea@cs.pub.ro

## ABSTRACT

This paper presents a preliminary study on how a deep neural architecture can be used to build a chatbot which simulates a conversation having the style of a given character, including famous personalities. Thus, we analyze how well a sequence to sequence (seq2seq) neural model can learn to generate meaningful responses for open domain dialogue and to imitate a persona. For the experiments, we trained the model to imitate two different personas: Romania's most famous poet, Mihai Eminescu, and Eric Cartman from the South Park series. The model was trained without using any handcrafted rules on an end-to-end dialog and is able to build new utterances to a wide number of user inputs using original distinguished expressions for both characters. However, the generated utterances still need to be improved for a meaningful dialogue.

## Author Keywords

Conversational agent, Dialogue systems, Word embeddings, Deep Neural Networks, Text style transfer

## ACM Classification Keywords

I.2.7 Natural Language Processing; H.5.m HCI, Miscellaneous

## INTRODUCTION

Conversational agents, also named chatbots when discussing about agents responding in a chat dialogue, have always been a topic of interest at the border of Natural Language Processing (NLP) and Computer-Human Interaction (CHI). They have been envisioned as an alternative, more natural way of interaction with a computer application using natural language and dialogue. Due to the recent advances in deep neural networks, chatbots have recently become again a hot topic, especially in NLP. Moreover, conversational agents have always been linked to the Turing Test, which is one of the most challenging problems in computer science research: can we train/build machines to act like humans in a conversation?

When developing a conversational agent there are several common challenges that sometimes are not so obvious. For example, retrieval-based models for chatbots use predefined responses and pick one of them using a set of heuristics, which can be either rule-based using logic and regular expressions or machine learning approaches for determining the best answer for a given input from a predefined list of answers. These models don't make any grammatical mistakes, their answers are syntactically

correct, but can't handle questions for whom they don't have a predefined response. One recent example for this type of system is Google's Smart Reply system recently developed for Gmail [1].

On the other hand, generative models can handle new cases because they don't rely on (any) predefined responses and generate their own response starting from the human utterance they have to respond to. These generative models are special because they can give the users the sensation that they are talking with a real human mainly due to the multitude and variety of possible replies. As there are no predefined answers, these models need to learn possible ways to build answers using a large collection of conversations. Thus they require large dialogue datasets for training and their flexibility is limited due to the fact that grammatical mistakes are more frequent, especially for long sentences, mainly because they do not use any underlying grammars when generating responses.

Many of the conversational agents built until recently have been aimed at solving specific tasks. However, at this moment there is a growing interest in developing open domain chatbots. In an open dialogue system, there isn't a well-defined purpose or intention for the conversation, therefore the user can drive the conversation on any topic he/she wants. This is closely related to Artificial General Intelligence [2] since the amount of knowledge and the number of topics needed in order to create an authentic conversation are practical infinite in this scenario.

At the opposite side, a close domain conversational agent is easier to develop using retrieval-based models since the agent has a predefined task, like booking a hotel room or buying flowers from a shop. In this case the number of possible inputs and outputs is limited and the system just needs to fulfill its specific purpose. In this case, the users are not expecting the chatbot to provide answers on any other topic than the one it was designed for.

The paper continues with a section that provides a brief motivation for our proposed approach for solving the inconsistency of a chatbot's utterances by incorporating a specific linguistic style. The following section discusses the most important recent works on building a persona-based conversational agent by inserting speaker embeddings into neural conversational models at training, but also by using standard rule-based approaches. Then we describe our proposed method, focusing on dataset construction, the

model, and the method used for training. The last two sections of the paper present preliminary results for the two case studies and concluding remarks.

### MOTIVATION

In this paper we focused on developing an open domain chatbot using a generative model which should be able to provide answers with a linguistic style similar to a given character. We call this type of conversational agent a model with personality. Although this term might be an overstatement for the current results, but it is an ambition for the future. We chose to implement a neural network architecture based on the sequence-to-sequence (seq2seq) model [3] for generating the response of the chatbot starting from other sequences of text, such as the human input. Seq2seq models have recently been very successful in other NLP tasks, such as machine translation and speech recognition. This model consists of two layers of recurrent neural networks using Long Short-Term Memory (LSTM) cells, an encoder and a decoder [3]. The input sequence is fed to the encoder which generates a word embedding for the entire input and for each word individually using a fixed size embedding vector. This vector is then passed to the decoder layer which generates the response one word at a time. To allow the decoder to have a more informed access to the input sequence, an attention mechanism was later introduced to the standard seq2seq model [4].

There is a major challenge when developing a chatbot using neural networks for two main reasons: the lack of a coherent personality and the inability of the network to generate longer sentences which are correct syntactically. In this paper, we aim to solve the first problem. When the model generates a response, we consider that it is desired that the response is related to a certain human persona. Most of the models generate plausible responses for the previous utterance of the user, but these responses are not consistent over a longer conversation, as can be seen in Table 1.

The major cause for this issue is that the training dataset for the neural network generative model (e.g. seq2seq) uses conversations from a multitude of different users in various contexts. On one hand, it is important to have a large dataset with conversations from various contexts in order to be able to respond to as many user inputs as possible. However, this creates inconsistency for the utterances generated by the chatbot.

We have tested our proposed approach in two completely different contexts. First, we developed a chatbot which should use the linguistic style of Romania's most famous poet, Mihai Eminescu. Then we tried to see the results on the character of Eric Cartman from the South Park series. In order to model the chatbot to talk in the same style as a given character, we started by collecting published works of Mihai Eminescu and movie transcripts from South Park. We first trained our neural model on a large dialogue corpus, then refined the training on the specific character

dialogues alone. We consider that after this process, the resulting embedding vectors incorporated some of the linguistic style of the character - and we call this a chatbot with a specific personality.

<b>User:</b> what is your name
<b>Bot:</b> pris
<b>User:</b> how old are you
<b>Bot:</b> eighteen
<b>User:</b> where do you live
<b>Bot:</b> in the closet
<b>User:</b> how many legs do you have
<b>Bot:</b> twenty eight thousand

**Table 1. Sample conversation with baseline chatbot**

### RELATED WORK

Until recently it was believed that a conversational agent must use a large number of predefined rules in order to handle all/most the cases in which it could respond to an input from the user. These rules were usually defined using specific markup languages and using regular expressions, such as Artificial Intelligence Markup Language (AIML) [5]. Chatbots using AIML, like ALICE Bot, that mapped a user input to a rule which was then used to retrieve a predefined response proved that this approach is practically impossible to develop an open-domain chatbot. The generated responses feel artificial, not relevant for many user inputs, and repetitive.

Changing the rule based approach with a machine learning approach was a step further. The Google Smart Reply [1] was trained on a corpus of labeled utterance responses in order to determine which utterance has a simple short response. The model encoded utterances and responses from the dataset into vectors using word embeddings and then measured their similarity by computing the dot product of the vectors. A large value proved that the vectors were similar and that corresponded with a positive label that the response was relevant for the given utterance. Even though these models can handle only conversations in a closed domain (e.g. meeting requests) due to the training corpora, they are very popular.

There have been several works on developing a chatbot with personality using generative neural networks, most noticeable of them is Li et al. [6]. They proposed to generate the embedding vector of a persona, based on the real person who generated the training dataset. This vector is called the speaker embedding. Afterwards, the speaker embedding was injected into the decoder of a generative model at every time step during training. This is equivalent

to forcing the model to remember information about a person when generating a response.

We used a different approach, starting from the work of Nguyen et al. [7] who propose to pretrain the word embeddings for the decoder using a large dataset and then to build character embeddings by further training the decoder embeddings on a smaller dataset containing only the responses of the embodied persona. We have used a similar approach in our work, both for Romanian and for English in two different domains: generating texts with the linguistic style of Romanian poet Mihai Eminescu and with the persona of the movie series character, Eric Cartman from South Park.

There have also been some works that focused on developing a conversational agent to provide responses about the life of a given historical personality [8, 9]. Although this may seem similar to the topic of this paper, it differs in the fact that the generated responses are not individualized with the linguistic style of each character. Rather, they generate responses based on a mix of topic based AIML rules extracted from Wikipedia pages [8] or using a combination of knowledge base text generation and retrieval-based models from Wikipedia [9]. Thus the generated utterances do not carry the linguistic style of the persona.

### PROPOSED SOLUTION

One major problem when training a generative neural model is underfitting caused by the small size of the training set. To overcome this problem, we first trained a generative chatbot on existent English large dialogue corpus. After this model was successful, we applied the same steps on training a chatbot on our own corpus gathered for Romanian.

### Dataset Construction

For English, we used the Cornell Movie-Dialogs Corpus by Danescu-Niculescu-Mizil and Lee [10]. As a short description, the corpus contains 220,579 conversational exchanges between 10,292 pairs of movie characters and involves 9,035 characters from 617 movies. The problem that appears when trying to embed a persona from this corpus is that conversations have different styles for each movie. This leads to an unrealistic persona for the chatbot which will provide inconsistent answers between turns. Thus, the model trained on this corpus generated plausible responses, but not consistent ones.

In order to add more authenticity in the conversation we added dialogs from the South Park movie series. We collected more than 70,000 lines of dialogue together with character information from the following dataset <https://www.kaggle.com/tovarischsukhov/southparklines>. Because each context has its own speaker now it is easy to extract lines for a certain character. This way we can make

a clear distinction in the dialog between different movie characters.

Cornell Movie Dialog Corpus (English)	South Park Corpus	Cartman Corpus
220,579	75,135	10,342
Subtitles from Movies Corpus (Romanian)	Eminescu Corpus	Eminescu letters Corpus
1,441,313	106,202	3,835

**Table 2. Number of dialogue pairs for the different corpora (English and Romanian)**

At the end, we gathered over 10,000 pairs of question-response for the character of Eric Cartman. Due to the small size of this dataset, we were not able to use these replies alone for training the neural generative model.

For the construction of the Romanian dialogue dataset, first we started by collecting movie subtitles in Romanian. At the end, our dataset contained 1,441,313 conversational replies. As conversations could not be assigned to a specific movie character due to the nature of the data, we considered one line as a given utterance and the following one as the reply to the previous one (this is the way subtitle lines appear in a movie subtitle file).

For the construction of the Mihai Eminescu corpus, we collected texts from the prose, journalism and personal letters of the author. We could not make a clear boundary between Eminescu's replies and ones of other characters as in the previous corpus, so we extracted only the replies from his personal letters and used them as utterance-response pairs. The size of this datasets are presented in Table 2, together with similar information about all the corpora used by our paper.

One important thing to notice for the Eminescu dataset is that the pairs of question-response in our corpus are not always aligned. For this reason, the topic for some utterances may differ from the topic of the corresponding answers. This can lead to generating responses that are not in the same context with the question.

### Model Description

We used a slightly modified seq2seq model from Tensorflow (<https://www.tensorflow.org/>). It is composed of an encoder and a decoder which can be viewed as two different language models. The encoder transforms the input sentence into an embedding vector. The decoder uses the embedding vector as a map and generates the response one word at a time.

The model used in the experiments has 3 layers composed of 256 Gated Recurrent Unit (GRU) cells. GRU cells have been preferred to the standard LSTM cells used by the



seq2seq model as they are a bit simpler and thus require less training data.

In order to improve the results, we reversed the input into the encoder and we used an attention mechanism to allow the decoder to have more informed access to the input sequence. The original decoder was implemented using beam search in a greedy fashion. As the result for each step individually was affected by local optima and the generated response was not very relevant, we decided to keep the best  $k$  candidates and perform a beam search thus exploring more results and getting an output which is somehow closer to the global optimum.

We used padding and bucketing to solve the problem of variable length of input sentences. For example, we consider buckets of sizes [(5, 10), (10, 15), (20, 25), (40, 50)]. For example, if the length of the input sentence is 4 and the length of the output sentence is 7 we put this conversation pair into bucket (5, 10) and pad with unused words until we fill to the maximum bucket size (e.g. 5 for encoder and 10 for decoder).

For the word embeddings, which is the method of representing words in a low dimensional vector space used by the seq2seq model, we used a size of 256 both for encoder and decoder.

### Training

In order to overcome the problem of the small datasets for a given character, we used a different method for training:

- In the first phase, we pretrained the seq2seq model on large conversational datasets for 45,000 steps for both characters (Eric Cartman in English and Mihai Eminescu in Romanian). The first dataset was composed from Cornell Movie Dialog and South Park datasets. The second one was composed from Romanian movie subtitles dataset and entire Mihai Eminescu datasets.
- In the second phase, we trained on the specific corpus from which we extracted the character lines for about 10,000 steps. Thus, we trained on the South Park dataset and on the entire Mihai Eminescu dataset. We consider that this step is important to fine-tune the word embeddings for the given context of the character.
- In the final phase, we concentrated our training only on the character specific lines for 5,000 steps. We used utterance-response pairs from Cartman corpus and created utterance-response pairs from Eminescu's personal letters. This last step actually embodies the character persona into the model, mainly by changing the language model of the decoder.

We used separate vocabularies for the encoder and the decoder. In order to create a vocabulary containing only the most frequent words, we plotted their frequency versus number of words and chose the vocabulary dimension that covers most of the area under the curve (see Figure 1). Any other word was replaced with a special token called *UNK*

(unknown). If the number of unknown tokens is large, the model learns to output UNK more frequently than other words and thus the conversation is flawed. We chose the vocabulary size to be 40,000 for both the English and the Romanian datasets.

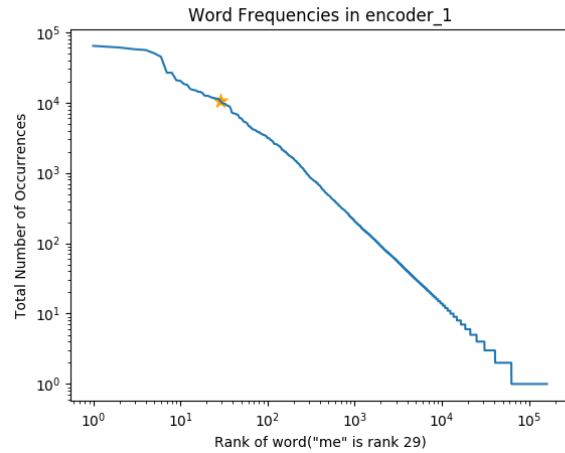


Figure 1. Word frequencies in the vocabulary of the encoder

### RESULTS

As previously mentioned, we have tested our approach on two characters: Eric Cartman in English and Mihai Eminescu in Romanian. The models were able to embed some of the linguistic style (personality) of the two characters. The two conversational agents speak in the specific style of each character as can be seen in Tables 3 and 4. As there are no proper metrics for the evaluation of conversational agents, we used both human evaluation and metrics used in machine translation by comparing the generated utterance with the one issued by the character in the dataset (given a specific input from the dataset).

<b>User:</b> what is your name
<b>Cartman-Bot:</b> ohhhh kyle ohhhh mom
<b>User:</b> who are you
<b>Cartman-Bot:</b> you have to be kidding me
<b>User:</b> where do you want to go
<b>Cartman-Bot:</b> what does that mean i will just have to go back to school
<b>User:</b> who is your best friend
<b>Cartman-Bot:</b> ohhhh my god they killed kenny

Table 3. Sample conversation with Eric Cartman-based chatbot

As a side-note, we trained the Eminescu chatbot encoder with sentences also used by the decoder. As a consequence,

instead of testing with normal sentences we had to use sentences that were similar with those from his letters.

If we compare the responses from the baseline chatbot with the responses from the Cartman chatbot we observe that the first ones are fitting better in the conversational context of the corresponding questions, but they lack the consistency of a personality. On the other hand, the responses from the Cartman chatbot do not always seem to be in the same context as the corresponding questions, but they incorporate the writing style of the character (see Table 3).

One possible reason might be that the baseline chatbot was trained on Cornell Movie Dialogue corpus which has better aligned pairs of question-response compared to our Cartman training dataset. This means that if we construct our dataset by supervising the correlation between questions and answer, we can improve the quality of the generated responses.

<b>User:</b> te sarut de mii de ori
<b>Eminescu-Bot:</b> al tau pentru totdeauna eminescu
<b>User:</b> ingerul meu iubit
<b>Eminescu-Bot:</b> UNK cu care eu n am fost niciodata
<b>User:</b> am suferit atat de mult pentru iubire
<b>Eminescu-Bot:</b> ma UNK ce as dori a ma ucide
<b>User:</b> ma pierd in negura timpului pierdut
<b>Eminescu-Bot:</b> din ceea ce nu voi sa se stie niciodata

**Table 4. Sample conversation with Mihai Eminescu-based chatbot**

Comparing Tables 3 and 4, we observe that Cartman responds to more general questions than Eminescu. This is because its encoder was trained on a more general dataset compared with the Romanian dataset for Eminescu. Moreover, the Eminescu chatbot generates more often UNK tokens than the Cartman one, showing that we need to have different vocabulary sizes for the two datasets.

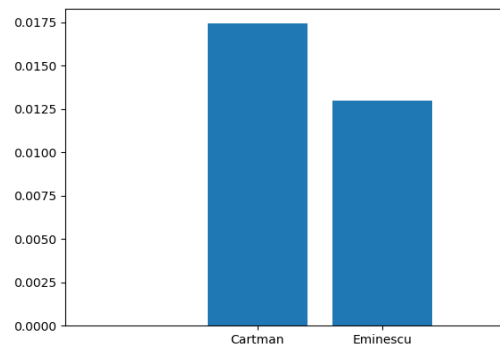
In the future, in order to add diversity we will have to build a more general corpus for Eminescu by adding a new corpus with general utterances and align them with our corpus. After we can train the encoder only on the new corpus added and the decoder on the Eminescu corpus. This way we can test the Eminescu chatbot using normal sentences and the decoder will generate sentences having the desired style of the famous Romanian author.

### Evaluation

The purpose of the performed evaluation is to measure that the conversational agent fulfill its task for a specific dialog. For open domain models there is no specific goal in the

conversation. In this case, metrics like BLEU score which are used for machine translation are not perfectly suited because they are based on word matching and they show no correlation with human judgments as shown in [11]. The BLEU score measures the number of n-grams that appear in both in the candidate and the target utterance and thus represents the similarity between them. A strong similarity is correlated with a BLEU score closer to 1 and a low similarity with a score closer to 0.

We measured the BLEU score between the chatbot output and the actual response from the datasets for a given context utterance. As can be seen in Figure 2, there were very few similarities between the generated response and the actual response from the dataset. This is the reason why when testing with BLEU the scores are very low.



**Figure 2. BLEU score for Cartman and Eminescu chatbots.**

As an alternative evaluation, we designed an empirical evaluation using 10 human subjects and we used it as a metric to rate our model. This metric showed a higher score than the previous BLEU metric for the generated responses. For testing we picked 20 utterances that were not in the training corpus and we labelled them with 2 responses: the actual response from the South Park movie series and the Eric Cartman chatbot's output. We scored the generated responses that were tagged as responses from the movie series. The model received a score rate of 15%.

The human testers provided us with qualitative insights of our model. We observed that the chatbot repeated certain words that were not fit to the context in most of the generated sentences. This is the reason why most testers picked those sentences as being generated by the chatbot. We also observed that when generating longer sentences the chatbot was prone to grammatical errors or generated sentences that were not coherent. We think that this is caused by the maximum likelihood estimation (MLE) which finds the values for the parameters of our model which maximize the likelihood of an observation given those parameters. That is the reason why the responses contained words which had a high probability, words that were prone to appear many times in the training datasets.

Although the score received from the human testers is better than BLEU, the results can still be improved. This motivates us to enhance our model to achieve a higher score with human evaluation in the future.

As a last note, the fact that we could not properly evaluate our model implies that we also could not optimize it. As a result we could not tell if our model could have been improved or not if we had increased the number of training steps or changed the parameters for the seq2seq model.

## CONCLUSIONS

This paper presented a method for building a chatbot using generative neural networks for determining the response in an open domain system that is able to answer using the writing style of a given human character. Thus, we have proposed to embed the personality (writing style, in this case) of the human character into a chatbot by modifying the training phase of the neural network. Slowly reducing the size of the datasets during the 3 phases of training by keeping only the character lines proved to be successful.

For future improvements we will concentrate on making the conversational agent able to remember information from the past utterances in the conversation by using an additional embedding for encoding recent past input sentences. More, we aim to improve the human evaluation score obtained by implementing a different architecture using Generative Adversarial Networks [12]. This architecture consist of two models: a generator (which is a seq2seq model) and a discriminator that classify utterances as human or machine generated. The discriminator is equivalent to the human testers in the Turing Test. The generator produces sentences and tries to fool the discriminator into believing that the sentences are generated by humans. The goal of this model is to generate utterances that are indistinguishable from human dialogue. Moreover, in this case we propose to use reinforcement learning to train the generator and the discriminator in a similar way to [13].

We will try to make a final improvement by adopting a Bayesian approach in the maximum likelihood framework as suggested in [14]. This way we will add two advantages: more expressiveness by an explicit representation of uncertainty and eliminating overfitting by regularization. This will also eliminate the problem with optimizing our model.

The adoption of a different training method in different phases was the key approach that gave a major breakthrough in personality of chatbots using neural networks. Future work in these directions will be developed.

## ACKNOWLEDGMENTS

This research was partially supported by University Politehnica of Bucharest through the Excellence Research Grants Program UPB–GEX 13/30.09.2016.

## REFERENCES

1. Kannan, A., Kurach, K., Ravi, S., Kaufmann, T., Tomkins, A., Miklos, B., ... & Ramavajjala, V. (2016). Smart reply: Automated response suggestion for email. *arXiv preprint arXiv:1606.04870*.
2. Goertzel, B. (2007). *Artificial general intelligence* (Vol. 2). C. Pennachin (Ed.). New York: Springer.
3. Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. In *Advances in neural information processing systems* (pp. 3104-3112).
4. Chorowski, J. K., Bahdanau, D., Serdyuk, D., Cho, K., & Bengio, Y. (2015). Attention-based models for speech recognition. In *Advances in Neural Information Processing Systems* (pp. 577-585).
5. Wallace, R. S. (2009). The anatomy of ALICE. *Parsing the Turing Test*, 181-210.
6. Li, J., Galley, M., Brockett, C., Spithourakis, G. P., Gao, J., & Dolan, B. (2016). A persona-based neural conversation model. *arXiv preprint arXiv:1603.06155*.
7. Nguyen, H., Morales, D., & Chin, T. (2016). A Neural Chatbot with Personality. Technical Report, *Stanford University*.
8. Haller, E., & Rebedea, T. (2013). Designing a chat-bot that simulates an historical figure. In *Control Systems and Computer Science (CSCS), 2013 19th International Conference on* (pp. 582-589). IEEE.
9. Bogatu, A., Rotarescu, D., Rebedea, T., & Ruseti, S. (2015). Conversational Agent that Models a Historical Personality. In *Romanian Conference on Computer-Human Interaction (RoCHI) 2015*. Matrix Rom (pp. 81-86).
10. Danescu-Niculescu-Mizil, C., & Lee, L. (2011). Chameleons in imagined conversations: A new approach to understanding coordination of linguistic style in dialogs. In *Proceedings of the 2nd Workshop on Cognitive Modeling and Computational Linguistics* (pp. 76-87). Association for Computational Linguistics.
11. Liu, C. W., Lowe, R., Serban, I. V., Noseworthy, M., Charlin, L., & Pineau, J. (2016). How NOT to evaluate your dialogue system: An empirical study of unsupervised evaluation metrics for dialogue response generation. *arXiv preprint arXiv:1603.08023*.
12. Li, J., Monroe, W., Shi, T., Ritter, A., & Jurafsky, D. (2017). Adversarial learning for neural dialogue generation. *arXiv preprint arXiv:1701.06547*.
13. Li, J., Monroe, W., Ritter, A., Galley, M., Gao, J., & Jurafsky, D. (2016). Deep reinforcement learning for dialogue generation. *arXiv preprint arXiv:1606.01541*.
14. Fortunato, M., Blundell, C., & Vinyals, O. (2017). Bayesian Recurrent Neural Networks. *arXiv preprint arXiv:1704.02798*.

# A pipe-based approach for service-oriented UIs

Mathias Kühn and Peter Forbrig

University of Rostock

Albert-Einstein-Str. 22

18051 Rostock, Germany

{mathias.kuehn, peter.forbrig}@uni-rostock.de

## ABSTRACT

Service-oriented architectures as instruments for distributed applications can be used in contexts that are grounded on networked devices. Specifications of web service interfaces can be instantiated across different platforms. Combining this method with specifications of platform-independent user interface models would allow to access applications from various interactive devices. According to this, application states and events need to be considered for individual consumers. Additionally, interpreters and generators need to be available for each consumer.

The paper proposes an approach for consumer-specific configurations of mentioned tools that are based on models derived from the Cameleon Reference Framework. In the end, pipes are configured that allow using web services by generated user interfaces. Configuring these pipes with models for layouts and dialogs is focus of the paper.

## Author Keywords

Model-based user interfaces, multi-path development, interactive system design

## ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

## INTRODUCTION

Software that is based on service-oriented architectures can be used together with various platforms. Web services can be addressed from consumers nearly with any platform. Related interfaces can be specified with Web Service Definition Language (WSDL) and need to be generated at runtime. Also user interfaces (UIs) can be specified with model-based languages that can be generated using methods which are grounded on the Cameleon Reference Framework. Combining both methods for generating interfaces would allow using functional cores together with UIs.

Generated user interfaces need to be flexible due to the fact that user abilities can vary a lot. UIs for users that are able to see differ from UIs for users that are blind. UIs can for instance be graphical or vocal. However, needed tool support could apply methods for generating UIs on different levels of abstraction. UI events need to be reified and abstracted at runtime. Additionally, domain data need to be communicated for any application. Especially web services are good to achieve this goal, because serialization & marshalling are supported by default.

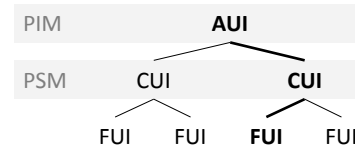


Figure 1. Transformation paths for generated FUIs.

Figure 1 shows a visualization of paths that generated final user interfaces (FUIs) can be transformed to when applying the Cameleon Reference Framework. Many FUIs can be generated from a specific concrete user interface (CUI), for instance for graphical UIs. Additionally, many CUIs can be transformed from specific abstract user interfaces (AUIs), for instance for graphical or vocal UIs. However, single FUIs have specific paths that they can be transformed to. These paths could be represented by pipes that are configured by models which describe UIs at different levels of abstraction. Applying pipes would allow transforming AUI to FUI and FUI to AUI whenever platform-independent models are needed. These models allow a broader range of applications for interactive systems.

Service-oriented architectures [9] are well-known in industry these days. Implementations are independent of specific platforms, what makes them more usable for a variety of different consumer platforms. Providing services allows using functions of software systems by any consumer. Different services can be combined to achieve new functionality. However, services also can be used to provide functional cores of interactive systems. One application core can be used by many different consumers. According to this, software architectures based on services would allow to use interactive systems together with any kind of UI.

The Cameleon Reference Framework [3] considers UIs on different levels of abstraction. Specifications can be made at the level of tasks, abstract, and concrete UIs. Additionally, UIs can be transformed to other levels of abstraction as well as to different contexts of use at the same level. Following that, the transformation types are translation, abstraction, and reification. In the end, final user interfaces can be generated from concrete UI specifications. These FUIs result from specific transformations along the different levels of abstraction. Configuring pipes for performing transformations could help in applying the framework.

The Arch model [1] separates interactive systems into different kinds of components that are depending on each other. The approach can also be seen as pipe, but it does not consider a parallel processing of data and events. Instead of this, it considers a sequence of components that assume certain relations between data and events. However, UI (presentation & interaction components) and data (domain adaptor & domain-specific component) also are considered on different levels of abstraction. Both kinds of components are related to each other on the level of dialog structure (dialogue component) that forces a sequential dependency.

The context of use [8] for interactive systems is an information space spanned by entities user (U), platform (P), and environment (E). It is defined as triple composed by U, P, and E. Relevant factors for users (U) are their physical abilities to see, hear, feel, smell, and taste. Relevant factors for platforms (P) are their potentials to externalize information visual (e.g. with displays), acoustical (e.g. with speakers), haptic, olfactory, and gustatory. For instance, displays can only be used by users that (temporary) are able to see. Additionally, relevant factors for platforms can be to internalize information with different kinds of sensors that are related to environmental factors. Relevant factors for environments (E) are light, sound, texture, aroma, and flavor. All factors are depending on information of space and time.

	Visual	Acoustical	Haptic	Olfactory	Gustatory
User	See	Hear	Feel	Smell	Taste
Platform (Ext.)	Display	Speakers	Actuator	Actuator	Actuator
Environment	Light	Sound	Texture	Aroma	Flavor
Platform (Int.)	Camera	Microphone	Sensor	Sensor	Sensor

**Table 1. Relevant factors for the context of use.**

Table 1 summarizes factors that are considered for the context of use. According to user-centered designs and as mentioned before, all stimuli that users physically can perceive with senses are listed here. Targeted platform components to internalize & externalize information as well as environmental effects, which surely can be other users, also are listed here.

#### RELATED WORK

UIs based on models are beneficial for targeting the generation of individual user interfaces that depend on specific contexts of use [6]. Model-based user interfaces (MBUIs) are independent of concrete platforms, but can be transformed so that they can be applied to specific exemplars.

Widgets of programming languages can be addressed from UI specifications that are based on these models. For instance, containers within MBUI specifications can be transformed to JFrame objects in Java Swing or to Window objects in C#. However, using model-based UIs allows addressing different platforms that can fit the variety of different end-user groups.

In [7] Paternò et al. propose an extension of WSDL that uses annotated specifications which can be combined with models for UIs. These annotations can be related to tasks of task model specifications in the end. By using this approach designers need to be aware of task model semantics and relations between tasks and web services. However, functionality of web services also can be used directly together with UIs that were generated from abstract models. These UIs are strongly coupled to specific functional cores.

In [2] Akiki et al. introduce the Cedar Architecture that uses web services for adapting UIs. Specifications of UIs also are communicated for making them context-specific then. Bindings to the application core are specified for individual clients and are not part of the UI model. Each consumer of functionality that is provided by web services need to adapt mappings to the application core by itself. However, UIs are context-specific and can be used by a variety of different groups of end-user clients.

The User Interface XML (UsiXML) [5] is one exemplar of model-based languages for UIs which is used for the proposed approach. UIs can be specified at the level of tasks, AUIs, and CUIs. Additionally, models for mappings, transformations, and contexts can be used for specifying dependencies to their application in general. Models for specifying behavior with state machine semantic is not covered by the language. However, UI layouts can be described in UsiXML and can be used for generating FUIs on any platform at runtime.

#### PROPOSED APPROACH

The approach in general considers data and events for communication based on web services at runtime. Domain-specific data need to be shared between service providers as well as consumers. Data continuously will be marshalled and de-marshalled for this reason. UI-specific data also need to be shared. UI-related events will consumer-sided be reified and abstracted. This is needed because only abstract events will be communicated between providers and consumers. Events additionally need to be mapped to functions that are implemented on provider side and also need to be mapped e.g. to widgets of generated FUIs.



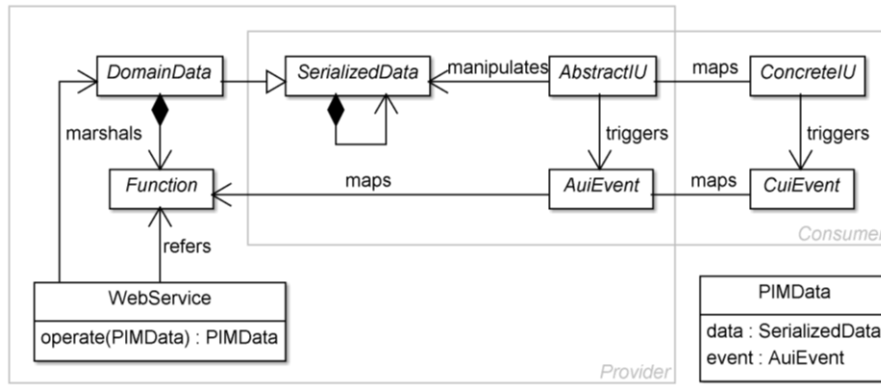


Figure 2. Architectural structure of the proposed approach.

Figure 2 shows a class diagram for the architectural structure of the proposed approach. The structure is separated into provider- and consumer-parts. Their intersection represents the structure which is used with web services. Class `WebService` uses instances of `PIMData` that are shared between consumers and providers. `SerializedData` only contains data and `DomainData` additionally contains functions. All instances of `Function` represent functional cores of applications. These functions are referred by web service instances. CUI models which consist of `ConcreteIU` instances are generated from AUI models which consist of `AbstractIU` instances. Instances of both are mapped to each other after generating UIs. Events that can be triggered later also are mapped to functions of the core.

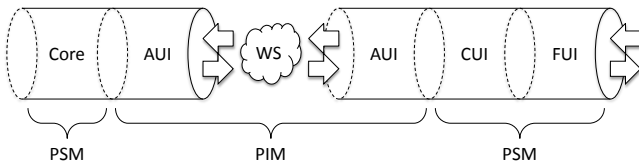


Figure 3. Visualization of the pipe for the proposed approach.

Figure 3 shows a visualization of the proposed approach for context-specific UI instances. Events of generated FUIs will be abstracted, communicated via web service interfaces (WS) together with relating domain data, and computed on provider side (Core), afterwards. Resulting events of the Core also will be communicated via the same WS interface together with domain data, reified on consumer side, and their effects will be presented on FUIs at last. Platform-specific models (PSMs) are used on provider and consumer side, but communication of events and domain data is made via platform-independent models (PIMs), what benefits the approach in general. Additionally, communication based on web services as well as serialized data are platform-independent, what makes the approach usable for any application of context-specific UIs with interactive systems.

However, transformations of AUIs to FUI instances also need to be specified depending on corresponding applications.

### Simple Example

This section introduces the approach at the example of a simple specification. The pipe will be configured by values that match the consumer-specific context of use for the UI.

```

<uiModel id="0">
  <uiModel id="01">
    <abstractIU id="011">
      <trigger id="0111"/>
      <!-- etc. -->
    </abstractIU>
  </uiModel>
  <uiModel id="02">
    <concreteIU id="021">
      <button id="0211"/>
      <!-- etc. -->
    </concreteIU>
  </uiModel>
  <uiModel id="03">
    <concreteIU id="031">
      <voiceCommand id="0311"/>
      <!-- etc. -->
    </concreteIU>
  </uiModel>
  <!-- etc. -->
</uiModel>

```

Figure 4. Example of a UsiXML specification.

Figure 4 shows a specification of a model-based UI that can be used for configuring pipes which allow reifications and abstractions of UI-specific events at runtime. A single AUI model specification (`id="011"`) can be transformed to CUI model specifications either for graphical (`id="021"`) or vocal (`id="031"`) UIs depending on the consumer-specific context of use. The FUI is generated from the CUI afterwards. However, events that are triggered on context-specific FUIs are transformed to their abstract representations in AUI for use with web services.

## CONCLUSION AND FUTURE WORK

The paper proposes an approach that allows using web services with context-specific UIs which individually will be generated for consumers. A pipe allows abstracting and reifying events that are relating to generated UIs. The pipe also supports serializing and marshalling data that need to be communicated between consumers and providers. It is assumed that functional cores of applications are provided by web services. These can be used by any UI that is generated. The pipe prepares events and data especially for consumer- or provider-side. Functions of the functional core only take events at the AUI level into account. However, the approach in general can only be used if sensors are available on consumer-side. It is assumed that end-user abilities can be detected by those sensors.

Next investigations need to be made on giving tool support for making specifications. The web service interface acts as façade [4] to the application core. It would be nice to generate this façade from an existing core. However, contexts of use are only considered when starting the application. It would be nice to react on changes at runtime so that generated UIs can always depend on the current context of use. For instance, a graphical user interface can be shown if end-users look at a display. If not, a vocal UI can be used for interacting with the application.

## REFERENCES

1. A metamodel for the runtime architecture of an interactive system: the UIMS tool developers workshop. In *ACM SIGCHI Bulletin*, vol. 24 no. 1 (1992), 32-37.
2. Akiki, P. A., Bandara, A. K., and Yu, Y. Engineering Adaptive Model-Driven User Interfaces. *IEEE Transactions on Software Engineering*, vol. 42 (2016), 1118-1147.
3. Calvary, G., Coutaz, J., Thevenin, D., Limbourg, Q., Bouillon, L., and Vanderdonckt, J. A unifying reference framework for multi-target interfaces. *Interacting with computers*, vol. 15 (2003), 289-308.
4. Gamma, E., Helm, R., Johnson, R., and Vlissides, J.: *Design Patterns: Elements of Reusable Object-Oriented Software*. Prentice Hall, 1st ed. (1994).
5. Limbourg, Q., Vanderdonckt, J., Michotte, B., Bouillon, L., and López-Jaquero, V. UsiXML: a language supporting multi-path development of user interfaces. In *Proc. of EHCI-DSVIS 2004*, Springer (2005), 200-220.
6. Paternò, F and Santoro, C. One model, many interfaces. In *Proc. of CADUI*, Springer (2002), 143-154.
7. Paternò, F., Santoro, C., and Spano, L. D. Engineering the authoring of usable service front ends. *Journal of Systems and Software*, vol. 84 (2011), 1806-1822.
8. The Context of Use:  
<http://www.w3.org/2005/Incubator/model-based-ui/XGR-mbui/#the-context-of-use>
9. Yuan, Y., Li, B., and Kreger, H. SOA Reference Architecture: Standards and Analysis. In *Proc. of SmartCom 2016*, Springer (2017), 469-476.

# Interactive Visualization of Spatial Multidimensional Data

**Bogdan Maxim**

Technical University of Cluj-Napoca  
Str. Memorandumului, Nr.28  
bogdy.maxim94@yahoo.com

**Dorian Gorgan**

Technical University of Cluj-Napoca  
Str. Memorandumului, Nr.28  
dorian.gorgan@cs.utcluj.ro

## ABSTRACT

The multidimensional spatial data understanding needs adaptive and performant algorithms executed on reasonable time. In order to discover the optimal complex algorithms, the visual analysis has to integrate the complex algorithmic processing over high performance processing infrastructures with interactive visual navigation throughout the data space. The visualization throughout the sparse value spaces is not a trivial task. This paper presents two interactive visualization techniques of the multidimensional spatial data, modeled by the Rasdaman technology. The paper presents the practical experiments on the two visualization techniques such as based on the intersection plane, and on the projection plane. The paper highlights and analysis the issues of the interactive visualization techniques throughout the sparse value space.

## Author Keywords

Visual analytics; sparse data; multidimensional spatial data; Rasdaman technology; visualization technique.

## INTRODUCTION

The great number of sensors (e.g. environment, satellite, etc.) and many other data sources from the domains of traffic and transportation, industry, agriculture, health, social life, and smart homes, produce huge quantities of data (Figure 1). While the capacity to collect and store new data rapidly grows, the ability to analyze these data volumes increases at much lower rates [10]. The analysis of these big data leads to new challenges in the data mining process, since analysts, decision makers, engineers, or emergency response teams depend on information hidden in the data. The visual analytics focuses on handling these massive, heterogeneous, and dynamic volumes of information by integrating human analytical reasoning by means of visual representations and interaction techniques in the analysis process.

The heterogeneous data are supplied by various sources such as computed data, sensor networks, simulated data, measured data in the field, or created and edited data. To store and process these data require very expensive storage and high performance processing resources. The main idea is not to just store data but analyze and valorize them. This a big challenge of transforming data in knowledge, as reduced and encoded form, and publishing data, as information [7]. The automatic data mining uses performant and adaptive algorithms that could be executed quite fast on the actual computing infrastructures (e.g. cloud, GPU clusters, grid, HPC architectures). Nevertheless, the creation, development and implementation of these algorithms needs creativity of the specialists by using flexible and performant tools.

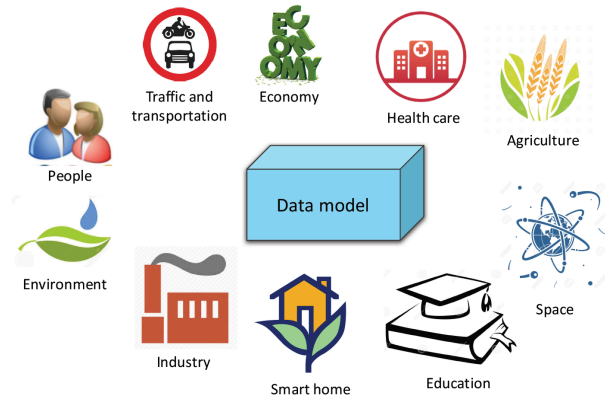


Figure 1. Multidimensional data sources.

One of such approach of exploring the data value space is the interactive navigation and visual analysis of the multidimensional data spaces. The specialists can identify by interactive visual navigation and analysis the particular values, critical regions, tendencies in static and dynamic data value space.

This paper presents the results of scientific research on exploring the interactive visual analysis on massive and multidimensional data, by using the high performance computation infrastructures. The paper highlights the advantages and the issues of such interactive navigation and visualization techniques.

The paper is structured by six sections. This section is Introduction and describes the domain, the challenge, and the objectives of the paper. The next section presents the related works reported in the scientific literature. The third section describes the Rasdaman data model. The fourth section describes the interactive navigation and visualization techniques. The paper follows by the experimental validation of the proposed visual analysis techniques. The last section concludes on the achievements and sketches the future works.

## RELATED WORKS

The visual analytics is an important subject in the scientific literature in the last decade. Keim et al. [10] provide an overview of visual analytics, its scope and concepts, addresses the most important research challenges and presents use cases from a wide variety of application scenarios.

Visualizations of multidimensional data are quite complex, and the interactive techniques support users to explore and select the presentation for better understanding [12]. Siirtola presents three conceptually different visualization techniques for multidimensional data: reorderable matrix visualization, enhancements to the user interface of parallel coordinate browsers, and a new visualization technique based on data glyphs and small multiple visualizations.

Sometimes multivariate data are not available as sets of equally multivariate tuples, but only as sets of projections into subspaces spanned by subsets of these attributes [16]. Xie et al. propose a visual analytics approach that helps the user understand the patterns and the uncertainties. Constraints that narrow the solution space can then be added by the user interactively during the iterative exploration.

Seo and Shneiderman offer a set of principles and a novel rank-by-feature framework that could enable users to better understand distributions in one (1D) or two dimensions (2D), and then discover relationships, clusters, gaps, outliers, and other features. Users of our framework can view graphical presentations (histograms, boxplots, and scatterplots), and then choose a feature detection criterion to rank 1D or 2D axis-parallel projections. By combining information visualization techniques (overview, coordination, and dynamic query) with summaries and statistical methods users can systematically examine the most important 1D and 2D axis-parallel projections.

In interactive data analysis processes, the dialogue between the human and the computer has to be not interrupted by slow computational mechanisms. Such integrated computations could be, for instance to reduce the dimensionality of the data or to perform clustering. Turkay et al. experiment in [15] the integration of such progressive methods within interactive analysis processes that involve high-dimensional data.

In [5] and [6] the authors presents the visual analytics techniques in exploring the solutions for strategic games.

In [11] Pahins et al. propose the notion of Hashedcubes, a data structure that enables real-time visual exploration of large datasets that improves the state of the art by virtue of its low memory requirements, low query latencies, and implementation simplicity. In some instances, Hashedcubes notably requires two orders of magnitude less space than recent data cube visualization proposals.

### SPATIAL MASSIVE DATA

The multidimensional space of data (Figure 2) is such a cloud that the user needs to analyze, process and discover their meaning. The data mining process integrated within the visualization loop needs optimal organization of the data model, rather than the classical relational database systems.

One such a solution of organizing the multidimensional spatial data is Rasdaman [1].

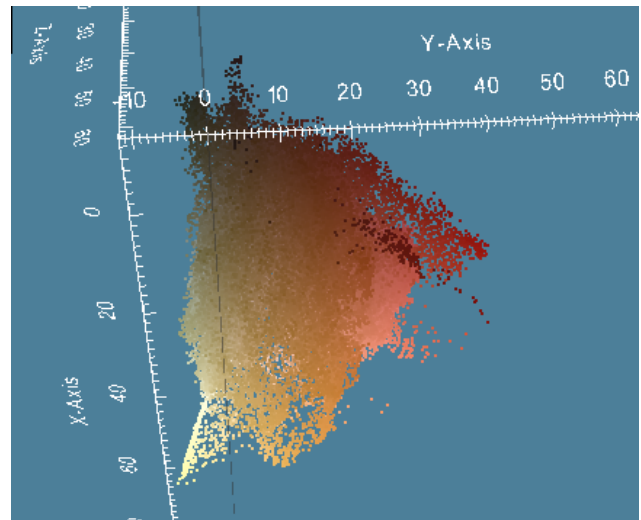


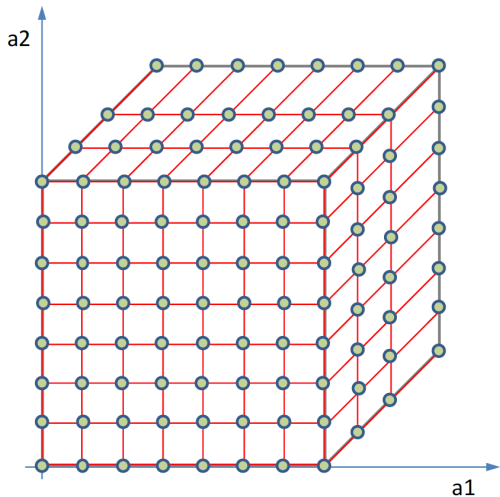
Figure 2. Value space of the multidimensional data.

### Rasdaman Data Model

Rasdaman is a domain-neutral array database system. It organizes data into a multidimensional array of homogenous, atomic elements, usually pixels or voxels. It has no limitation on the number of dimensions. It can serve, for example, for 1D measurement data, 2D images (like: satellite, tomographic, typographic, cartographic, infrared images), 3D x/y/t time series and x/y/z exploration space, 4D ocean and climate data, and can even exceed spatio-temporal dimensions [1]. The array managed by Rasdaman is Big Data, often single objects ranging into Terabyte and soon, into Petabyte sizes. Usually, this size is achieved by Earth and space observations.

The main concept on which Rasdaman arrays are based on, is called tiling. When an array is inserted into Rasdaman, it is first subdivided into smaller sub-arrays, called tiles, and then the tiles are being stored in a conventional database management system. Internally, Rasdaman uses PostgreSQL as its primary database management system. The tile is the main storage unit. Before inserting the tiles into the database, the system assigns them an index. This index is useful for the later operations which will be executed on them. Different tiling strategies (like: regular tiling, caveat, aligned, directional, area of interest, statistic and user generated) can be used when inserting data. Rasdaman provides a way to specify the index type used for tiling. It provides a general-purpose index, a directory index specific for interval objects and a special index which works with regular tiling. In its core, Rasdaman stores both the bulk array and the auxiliary geo-semantics for the real-world mapping of the arrays [4], [14].

Two special ways of storing the data in Rasdaman are: RectifiedGridCoverage and ReferenceableGridCoverage. The term "coverage" refers to any data representation that assigns the values directly to the spatial position. A coverage



**Figure 3. Multidimensional data space as a sparse volume.**

is a function which from a spatial, temporal or spatiotemporal domain to an attribute range.

The RectifiedGridCoverage is a coverage which geometry is represented by a rectified grid. A rectified grid is a grid for which the function which transforms the grid coordinates is affine.

The ReferenceableGridCoverage is a coverage of which geometry is represented by a referenceable grid. A referenceable grid is a grid for which the function that transforms the grid coordinates are not affine (Rasdaman doc).

### Multidimensional data model

As the first important step of data organization, Rasdaman permits defining the data type which is going to be used. Each instance within the database is described by its type. A Rasdaman schema contains three categories: cell type definitions (which can be base types like floats or struct types like  $r/g/b$ ), multiple dimensional data type definition (which can be arrays over a base type and with some spatial domain) and collection type definitions which describe sets over some multiple dimensional data. Rasdaman supports adding new data types at runtime [3].

The basic unit of storage in Rasdaman is the cell. It may hold from a basic value, like integer, or float, to a much more advanced data structure, like  $r/g/b$  structs and anything that a user may think of. In general, any struct which is available in C/C++ can also be used in Rasdaman type definition, with the except of pointers and arrays. Nested structures are also supported.

By grouping multiple cells together, a more advanced data model is constructed. This is the multidimensional data array. Rasdaman allows the user to specify the cell type and the domain for the array. The domain can vary, and even boundless arrays are allowed. It is discrete and only integer

values are allowed (Figure 3). Rasdaman allows unlimited number of dimensions for the arrays [4].

Over multiple typed arrays, collections are being build. Collections are data types which contains two specific columns: a system-maintained object identifier and the array itself. Using this type of data organization, Rasdaman is able to handle such data in the relational way: foreign keys in conventional tables allow to reference particular array objects [2].

By considering time as one of the dimension, Rasdaman can easily take a series of images and store them into its internal database management system. The user defines the coverage for the collection, and the system is inserting the images based on those properties. The user is also able to specify the time increment for the series. By using the coverage, Rasdaman can easily handle multidimensional geographical or satellite data which comes in the form of time series images [4].

### Data model interrogation

By using a relational model, the main data interrogation is being achieved using a query language. The Rasdaman query language embeds itself into the standard SQL and it is set-oriented processing. Besides the classic CRUD operations, the Rasdaman query language adds a new set of operations for the multidimensional data arrays. These operations are based on a minimal set of algebraically defined core. This set provides an array constructor (which establish a new array and fills it with values) and an array condenser (which provides scalar summary information from an array, like the sum, average, minimum, maximum, some and all cells).

Queries are parsed, optimized and executed in the Rasdaman server. Parallelism execution is also implemented. Different threads will process different parts of the array and the final result is assembled in the main thread.

Whenever a query is processed, the tiles of the selected array are processed sequentially, giving the possibility to query data with a dimension far bigger than the server memory.

When using the coverages, the user is able to manipulate data in an efficient way. Internal queries are optimized for the coverage section, giving a boost of performance for both the classic operations, and for some special ones. The user can select the domain of the dimensions which will be processed (both time and spatial) and a specific portion of the image to be displayed. The user is able to manipulate the images by zooming in and out, rotate the images and even process the images like addition, subtraction, multiplication and even user defined functions of the images. Those queries can be executed using the array constructor query which is a general-purpose operator. For a better performance, specific queries should be used when working with the coverages. Those queries include: for operator and array condensers.

By having a data model like this, the availability of the relational model is achieved. The information can be





able to select the three dimensions which he wants to visualize.

The section plane is defined at the camera position. The camera look direction is the normal to the plane. By moving the camera around, the user is able to inspect the different layers of the object.

The intersection plane is achieved by using the array constructor of the Rasdaman query language. Having the camera position and the normal to the plane, we compute the direction which is going to be used to make the cut on the x and z axis. We then compute the starting and ending point, based on the camera frustum. Having the starting and the ending point, we apply the Bresenham's line algorithm on the plane between those points. For every x and z points, we select all the y positions and add them to the array, as a column. After this, we display the newly created array to the user (Figure 6).

For the special cases of vertical and horizontal plane, we are using the select statement of the Rasdaman query language. The select statement lets us specify which dimension is constant and the domain of the other dimensions, and it will automatically build the new two-dimensional arrays based on our selection.

The newly created array is going to be send to the client and to be displayed there.

The main advantage of this technique is the presentation of the layers from which the objects are made of. By moving around, the user is able to identify different "extreme points" of the data, which allows into a better understanding of the composition of the object. Also, the user can identify different abnormalities which can be present. The user can also specify time as one of the dimensions. By doing so, the user may follow the abnormality and see its evolution, and, in some cases, it can even predict what it is going to happen there.

By knowing all the layers from which an object is made, the user can specify the layer's visibility. This is extremely useful in the case of projection plane technique. In that case, the user is able to see a different point of view for the abnormality and deduct more results about its origin and evolution. By specifying the visibility, the user can see different layers of the object in a projection plane, making different combinations of layer's visibility to achieve a much more complete view over the model.

#### *Projection plane techniques*

The second technique that we propose is the projection plane technique. This allows the user to visualize the space in a 3D manner. The user is able to select which dimensions he wants to visualize, the resolution of the image and the camera parameters. He is also able to manipulate the camera parameters during the runtime of the visualization. Using this method, the user gains a better understanding of the outside view of the object and the way the object resides in the world.

The user selects the three dimensions he wants to visualize them.

The main technique implemented here is the Ray-casting method. Given the projection method which we want to use, we are sending rays from the user's camera position into the world, based on the look vector. At runtime, we compute the frustum coordinates in the world [9].

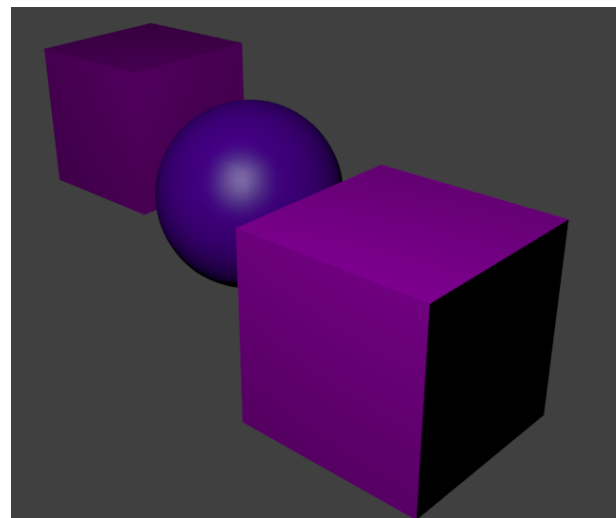
The rays compute the visibility of the layers as they pass on and will give the final color of the given pixel. Here, the resolution of the image is the most important parameter. Given a smaller resolution than the space, only fragments of the object will be seen. This is resolved by triangulating (Delaunay triangulation), the points returned by the ray-caster and interpolating the triangles (triangle rasterization). If the resolution is the same, no interpolation will be executed, as the rays hit all the visible point of the object. If the resolution of the image is bigger than the resolution of the space, more rays will hit the same points of the surfaces (Figure 7).

When the rays are send from the camera position to the objects, they compute the final pixel value based on its visibility and the other points' visibility. It is used the iterative visualization formula for multiple transparency layers:

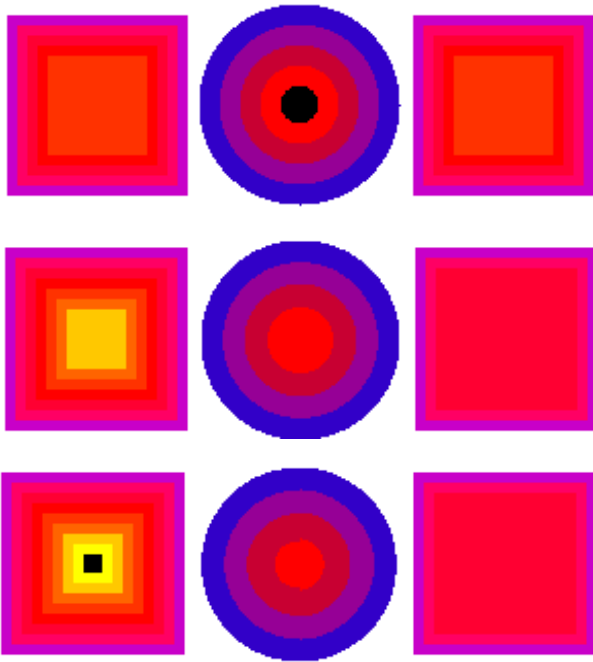
$$C_0 = \alpha_1 C_1 + \alpha_2 (1 - \alpha_1) C_2 + \alpha_3 (1 - \alpha_1)(1 - \alpha_2) C_3 + \alpha_4 (1 - \alpha_1) (1 - \alpha_2) (1 - \alpha_3) C_4 + \dots$$

Where C represents the layer color, and alpha represents the layer's visibility.

By doing so, different effects can be achieved. Some layers can be fully transparent, some layers can be partially transparent, and some layers can be fully opaque. This is extremely useful in the case where the user wants to see exactly a layer, without being distracted by other. The user is



**Figure 5. Scene of objects stored within the rasdaman data model.**



**Figure 6. Interactive visualization by intersection plane: (top) vertical plane cuts through three objects - two cubes and one sphere. The objects are made from different colored layers; (middle) oblique plane; (bottom) closer oblique plane.**

able at runtime to select which layers should be visible and which one should not. The background is always a transparent layer. The ray stops whenever the next colors have no influence over the final pixel color (for example, when it hits an opaque layer).

The main advantage of this technique is the visualization of the object from an outside point of view using the visibility method. In this way, the user which has identified an anomaly in the object, is able to see the way it looks from outside and gain different information which, combined with the section plane visualizing, gives a whole new perspective over the problem. The visibility mechanism is extremely useful when the user needs to be focused on a specifically part from the object and does not want other layers to distract them. This is also useful to see what lays after an opaque object and also to see what happens if a given object is transparent, rather than opaque. Different optical tests can be developed this way.

Another advantage of this method is letting the computer to search through the space, based on a given animation of the camera, and visualize different results returned by this. The user is able to pause the animation at any time, and also guide the computer to follow specific paths. In this way, useless search paths (and computations) can be avoided just by the user's intervention. This automation makes it easier for the user too, as he doesn't have to search by hand all the possibilities. His only intervention is wherever he wants to

deviate the normal search path as the old one can lead to nowhere.

### EXPERIMENTAL VALIDATION

Practical experiments have been defined and developed in order to prove the efficiency of interactive visualization of the sparse multidimensional data model.

The data model consists of spatial definition of a scene of 3D objects consisting of two cubes and one sphere (Figure 5). All objects are defined as concentric layers of various materials. The material of each layer is defined by its particular color and transparency.

The 3D scene of objects have been discretized and modeled in Rasdaman as multidimensional data with five attributes (i.e. x, y, z coordinates, color, and transparency). The discrete values generate a sparse description of the space. In order to visualize the objects, on the screen by high graphics quality at various high resolution, the system interpolates the screen space. It is not the scope of this paper to detail the discretization and interpolation of the objects. This paper concerns with interactive navigation and visualization by section and projection planes.

In such a manner the objects are viewed obviously by their exterior material. To visualize the inside objects, the systems offers two techniques: (1) by intersecting the objects through a section plane, and (2) by penetrating visually through outer layers to visualize the inner layers.

### Intersection plane use cases

Figure 6 presents the visualization cases by intersecting the scene of three objects by one vertical plane. In Figure 6.top the intersection plane passes by outer layers of the cubes and the middle of the sphere. Therefore, three layers are visible throughout the cubes and five layers throughout the sphere.

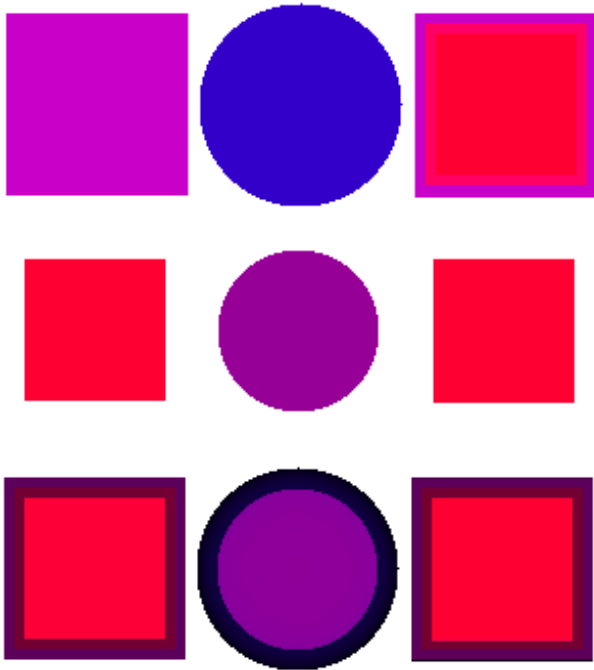
In Figure 6.middle the intersection plane is moved deeper in the left cube, and in In Figure 6.bottom the intersection plane passes through the center of the left cube.

By this technique are visualized the values along an intersection plane. The technique is less appropriate for volumetric visualization (e.g. cloud of values). For instance, instead of viewing by this technique the anatomical femoral bone by section, it is viewed as individualized 3D object by the projection plane technique.

### Projection plane use cases

Figure 7 exemplifies the visualization on the projection plane. Figure 7.top considers the outer space as a transparent space, and the three objects are viewed from outside.

Figure 7.middle presents the visualization of the inner layers of the three objects, by penetrating visually the outer layers. The visual penetration is controlled interactively by the user. The user can specify the viewing parameters, camera position, and the visibility of the outer space and object layers.



**Figure 7. Interactive visualization by controlling the visual penetration: (top) outside the objects; (middle) to inner levels; (bottom) 70% is the level transparency. Two layers are transparent for the cubes, and one layer for the sphere.**

Figure 7.bottom considers 70% the level transparency for two outer layers of the cubes, and one outer layer of the sphere.

By these two visualization techniques the user can navigate throughout the value space. By setting the viewing parameters (e.g. camera position, orientation, distance, zoom in and out, etc.) the user can move around the value space, and visualize the two cubes and the sphere from any position.

By the second technique, the user can adapt the transparency in order to visually penetrate the value space throughout soft values till the hard values. For instance, penetrate the skin, and flash of the human body, in order to visualize the bones.

By the projection plane technique, the user can identify continuous objects from sparse data. For instance, even the temperature of the atmosphere, have been provided just for some measurement positions, the system is able to visualize the same temperature spaces as contiguous sub-spaces (e.g. isothermal clouds).

By controlling the viewing penetration, the user could identify critical points, convergence, tendencies, surfaces, thresholds, and hard edges within the value space. The user could identify the limits of the homogeneous domains.

The main utility is in the cases when the value space in Rasdaman is filled in with data from the sensors, as sparse positions and various moments. The visualization techniques

support the visualization of the multidimensional space as a contiguous value space.

## CONCLUSIONS

The scientific research and the practical experiments have proved the high performance computation can be integrated within the visualization loop. More research and experiments will be developed in the domains of Earth Science, Earth Observation, medicine, climate change, and hydrological models.

The research and the presentation through this paper have proved the massive data can be understood by visual analytics techniques. This paper concerned mainly with two techniques of interactive visualization, which provide user with better access to data value space: intersection plane and projection plane.

New processing and visualization use cases will be analyzed and experimented in order to better integrate the processing algorithms on high performance computation infrastructures, and visualization platforms.

The Rasdaman technology will be experimented with different heterogeneous and sparse environmental data.

## ACKNOWLEDGMENTS

This research has been carried out in the Computer Graphics and Interactive Systems Laboratory (CGIS) of the Computer Science Department, in The Technical University of Cluj-Napoca.

## REFERENCES

1. Baumann, Peter. Management of multidimensional discrete data, *International Journal on Very Large Data Bases 3.4* (1994): 401- 444.
2. Baumann, Peter. *A database array algebra for spatio-temporal data and beyond*. International Workshop on Next Generation Information Technologies and Systems. Springer Berlin Heidelberg, 1999.
3. Baumann, P. Query language guide. URL: [http://rasdaman.eecs.jacobsuniversity.de/trac/rasdaman/browser/manuals\\_and\\_examples/manuals/pdf/ql-guide.pdf](http://rasdaman.eecs.jacobsuniversity.de/trac/rasdaman/browser/manuals_and_examples/manuals/pdf/ql-guide.pdf) (2013).
4. Campalani, Piero, Xinghua Guo, and Peter Baumann. *Spatio-Temporal Big Data*.
5. Catana M.C, Gorgan D., Analyzing Computer Game Strategies through Visual Techniques. *Proceedings of the 12th Romanian Conference on Human-Computer Interaction*, ISSN 2344-1690, pp.41-48, (2015).
6. Gorgan D., Catana M.C., Stefanut T., Visual analytics on multidimensional massive data (Tehnici vizuale de analiza a datelor masive multidimensionale). *Revista Romana de Interactiune Om-Calculator, Vol.8(3)* (2015), 237-255.
7. Gorgan D., Visual analysis on multidimensional massive data. *Journal of Aeronaut Aerospace*

- Engineering 2017, Vol.6(2)*, ISSN: 2168-9792 JAAE, (2017).
8. Heer J., Shneiderman B., Interactive dynamics for visual analysis. *Communications of the ACM, Vol.55(4)*, pp.45-54, (2012).
  9. Hughes, John F., et al. *Computer graphics: principles and practice*. Pearson Education, 2014.
  10. Keim D.A., Mansmann F., Schneidewind J., Thomas J., and Ziegler H., Visual Analytics: Scope and Challenges, chapter in *Visual Data Mining, Lecture Notes in Computer Science, Vol. 4404*, pp.76-90, (2008).
  11. Pahins C.A.L., Stephens S.A., Scheidegger C., Comba J.L.D., Hashedcubes: Simple, Low Memory, Real-Time Visual Exploration of Big Data. *IEEE Transactions on Visualization and Computer Graphics Vol. 23(1)*, pp.671-680, (2017).
  12. Siirtola H., *Interactive Visualization of Multidimensional Data*. Academic dissertation in interactive technology at the University of Tampere, Finland, (2007).
  13. Seo, J., Shneiderman, B., A rank-by-feature framework for interactive exploration of multidimensional data. *Information visualization, 4(2)*, 96-113, (2005).
  14. Tiling and Indexes in rasdaman, <http://rasdaman.org/wiki/Tiling> (last visit 2017).
  15. Turkay C., Kaya E., Balcisoy S., Hauser H., Designing Progressive and Interactive Analytics Processes for High-Dimensional Data Analysis. *IEEE Transactions on Visualization and Computer Graphics Vol. 23(1)*, pp.131-140, (2017).
  16. Xie C., Zhong W., and Mueller K., A Visual Analytics Approach for Categorical Joint Distribution Reconstruction from Marginal Projections. *IEEE Transactions on Visualization and Computer Graphics Vol.23(1)*, pp.51-60, (2017).



# SpiMO - Sitting Posture Monitoring System

## **Julia Birsan**

Faculty of Engineering in  
Foreign Languages, University  
POLITEHNICA of Bucharest  
Splaiul Independenței 313,  
București  
julliabirsan@gmail.com

## **Diana Stavarache**

Faculty of Engineering in  
Foreign Languages, University  
POLITEHNICA of Bucharest  
Splaiul Independenței 313,  
București  
diana.stavarache@gmail.com

## **Maria-Iuliana Dascalu**

Faculty of Engineering in  
Foreign Languages, University  
POLITEHNICA of Bucharest  
Splaiul Independenței 313,  
București  
maria.dascalu@upb.ro

## **Alin Moldoveanu**

Faculty of Automatic Control  
and Computers, University  
POLITEHNICA of Bucharest  
Splaiul Independenței 313,  
Bucharest  
alin.moldoveanu@cs.pub.ro

### **ABSTRACT**

Back pain is among the most common causes of missing work and contributes to about 93 million lost workdays and \$5 billion in health care costs every year. Sitting in an office chair for long periods of time can create lower back pain and can worsen problems that already exist. There are a number of factors that affect back pain, including level of movement, schedule, and desk set-up. Posture is a key factor that can induce pain over time; this includes sitting in the same position for too long or leaning too close to the screen. One may be able to avoid back pain or prevent its recurrence by improving the physical condition and learning and practicing proper body mechanics. We propose an efficient and time-effective tool to support the maintenance of a healthy position - a smart pillow. This pillow will track one's sitting posture on the chair, provides alerts when you don't have a right posture, provide you tips for exercises based on your posture, allow you to set up goals regarding your position and remind you when to take a break and sit-up.

The current paper offers development details related to the proposed innovative product, both from the hardware point of view (the pillow with sensors) and from the software point of view (the mobile application gathering signals from the pillow and providing learning alerts to the user), in the context of other similar applications available on the market and useful in learning how to have a healthier lifestyle.

### **Author Keywords**

healthy lifestyle; smart pillow; body posture; prevention

### **ACM Classification Keywords**

H.5.m.

### **General Terms**

Design

### **INTRODUCTION**

Posture is very important both at home and on the job especially if your jobs requires lots of hours of sitting at the desk. A back-friendly posture will prevent one from getting back and neck pains. Not only office chair sitting posture, but also driving posture to and from work can have a bad influence on one's health if the body position is ignored.

An important tool to support and maintain a good posture while sitting could be an ergonomic office chair. However, simply owning one is not enough - it is also necessary to adjust the chair to the proportions of one's physique in order to reduce the aggravation to the spine and improve personal comfort. There is no single type of office chair that is optimal for all patients, and people should determine their individual preference for comfort while following some simple steps of maintaining a proper posture.

The above mentioned are the main reasons we decided to develop SpiMo, a posture monitoring system which tracks the moves of the body. In this study, we investigate the use of mobile sensing technology to monitor spine stress in real-time during daily activity. [1]

Our proposed solution provides feedback with the use of a phone application and real-time notifications. SpiMo is a suitable solution for your spine health since this monitoring system could fit everybody's needs unlike the traditional ergonomic chairs.

The first part of the article consists in presenting the importance of a good posture to one's health, followed by some similar applications and the general, functional and technical description of SpiMo.

## POSTURE MONITORING SYSTEM

The use of sensor technology in monitoring the spine position and health in real-time has been rather limited. However there are similar solutions based on sensors or Kinect. One example would be Lumo Lift.

**Lumo Lift** [5] is a fitness wearable able to track your posture. How does it work? By resting against your décolletage, using algorithms and hardware sensors in order to measure the body's alignment. Every time you put Lumo on, you need to calibrate it and you also tell Lumo what posture you want to try to maintain (whenever it is slightly slouchy, or upright).

After that, during the day, Lumo tracks how often you stay in that right posture and when you start deviating from it. There is also a smartphone application, in which you can see your progress, but also the step count and distance.

The Lumo uses magnets in order to stick it on you, instead of clipped to clothes, so you actually wear it with a magnet. All the components are inside a polycarbonate tracker, while the plate holds the device in place on your clothes. The tracker acts like a button that you press to use it and program.

The device is smaller, more or less like the size of two quarters side by side, and just some centimeters thick. The magnetic configuration offers the user the possibility to personalize it, which most of the fitness trackers doesn't give. The user can create his own magnet clasp by using a strong one and, also decorate it.

Lumo is charging via a magnetic cradle with 2 pins that connects to the tracker in order to charge it. A full charge is about 2 hours and lasts five days. Unfortunately, Lumo does not have a LED display, but has a LED charging light.

The device is meant to be worn every day, in order to track your steps and posture. You simply attach it on your clothes, finding the right spot and, keep the posture you want. Calibrate the device by pressing twice on Lumo and in this way will enter in a Align mode. The device will vibrate in order to tell you that is recording your desired posture. The Lumo can be recalibrated as often as the user wants.

Once it is calibrated, the device tracks your posture and your motion all day long. Lumo has also a coaching mode that vibrates when you slip of your desired posture. The device has inside a microcontroller, a rechargeable battery and accelerometer, gyroscope and a Wi-Fi chip, used to send all data to a smartphone. With all these components, Lumo is raising the quality of life of many users, not only by making them more confident because of the straight posture, but also because can prevent users from developing back pains or other diseases caused by a bad posture.

Another example of spine monitoring system is **eCushion** which is designed for both clinical use as well as daily home use. A smart phone application gives real-time feedback on the distribution of sitting pressure on the cushion.

The implementation of eCushion consists of a textile sensor array, signal sensing and transferring unit and a computation and display terminal. A 16 by 16 textile sensor array is used to monitor the pressure distribution on cushion when users are sitting on it. The footprint of the sensor is 5/8 inch by 5/8 inch, and adjacent space is 1/8 inch. An Arduino based unit is used to sample and transfer the sensing data via bluetooth protocol. For sitting posture monitoring application, the sample rate is set as 10Hz. [7]

## SPIMO – SITTING POSTURE MONITORING SYSTEM

A good posture means your bones are properly aligned and your muscles, joints and ligaments can work properly. Poor sitting postures are common among adolescents as well as employees who work for prolonged hours. Spinal injuries are second only to the common cold as a cause of absence from work, with many of these problems emerging from poor posture habits. Since more spinal problems will inevitably lead to higher health costs and lower productivity, helping people maintain healthy spinal habits and reduce spine stress during daily activity is of considerable benefit. Some studies have showed that the some common sitting postures can lead to lumbar flexion and higher compressive forces in lumbar joints [2, 3].

Trying to avoid the adverse effects of bad sitting postures, the real-time monitoring of sitting posture received a particular attention and was used as a promising way in recent years. Increasing person awareness of bad posture means that the person can use her own back muscles to correct the spine, instead of using external support devices which could cause physical and psychological discomfort [4].

You may be able to avoid back pain or prevent its recurrence by improving your physical condition and learning and practicing proper body mechanics, but sometimes you do not have time for exercises. This is why Smart Pillow comes to help you. Smart Pillow is a smart pillow that will track your sitting posture on the chair, alerts you when you don't have a right posture, provides you tips for exercises based on your posture, allows you to set up goals regarding your position and reminds you when to take a break and sit-up.

### General Description

SpiMo is a smart pillow designed for both clinical and daily use at home, where the user can be remotely monitored. The system monitors sitting postures and provides real time feedback to the user for poor sitting posture but also monitors increased pressure on the back due to undesirable positions such as standing or sitting in a fixed position for a long period of time.

The system measures the weight of the user and detect which would be the right sitting posture. Furthermore, a daily summary report is automatically generated and it provides an account of the amount of time during the day that the user is sitting, to enable her to track her progress. The report also provides daily information related to posture angle and severity of poor posture. Data from the

sensors is acquired and transmitted to a database using a wireless module.

SpiMo is also able to suggest different type of exercises, for improving your health condition, based on user's height, weight, age, genre and report of the sitting postures. Basically, SpiMo is a virtual coach that analyzes your sitting posture on the chair and suggests you exercises for your health condition, but also sends a gentle vibration and sound alarm when imbalance is detected or you have sit in same position for a long period of time.

The feedback for the user must be presented in a smarter way. This is done in two ways, implicitly or explicitly. In the explicit way, the user is notified (for example with a pop-up message on his computer or mobile phone) of an incorrect posture, urging him to adjust his position accordingly. In the implicit way, the system stimulates the user in using a correct posture by vibration and sound feedback.

SpiMo is able to learn which are the good postures of a person and with an algorithm predicts your future sitting postures. For the cases when the user is spending more time than normal sitting, SpiMo uses the pomodoro technique [6] for the alerts: after each 50 minutes spent on the chair, the user should take 10 minutes break and sit up, doing some exercises or simply walking from a room to another.

### Functionalities

Inside the pillow are some thin sensors able to detect your sitting position. An Arduino Uno board was used in order to read the data from the sensors and via Bluetooth, sends your status to the mobile application. In this way, you will get notification whenever your posture is good or bad. Smart Pillow application is not only notifying you regarding the sitting posture, but also offers you some other features:

- **User profile** module allows you to personalize the application with your details (name, email, weight, height, date of birth, genre).
- Set **Goals module** allows you to set your good siting goals, in terms of hours, minutes.
- **Track posture:** one can see in real time whenever your posture is good or bad.
- **Track progress:** the Smart Pillow app offers your statistics regarding your posture, in order to know what would you have to do to correct your posture.
- **Exercises:** based on your profile information(age, height, weight and genre) and your posture progress, Smart Pillow offers you tips and tricks, exercises of how to improve your posture.
- **Coach ON/OFF:** with your coach ON, you will always receive alerts, tips and tricks and exercises regarding your posture. Smart Pillow coach uses pomodoro technique, that will also ask you to take a break and sit up at each 50 minutes. Smart Pillow together with

Pomodoro Technique will help you increase your work productivity.

### Technological description

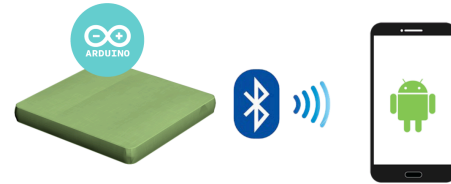


Figure 1, Smart pillow architecture

Smart Pillow improves your postures, because sitting too much on the chair and also in a bad posture, increases the chances to develop different diseases (type 2 diabetes, cardiovascular diseases).

The pillow is collecting data from the pressure sensors that are further establishing if the sitting position of the person is good or bad. The person is receiving a feedback directly on the phone, as long as the pillow is connected via Bluetooth to the mobile.

The overall system architecture is illustrated in Figure 1. The method of detecting postures includes pressure sensors that are connected to an Arduino board that computes the pressure distribution of the sensors, in order to establish if the sitting position is good or poor. All the data is sent to a database via Bluetooth and an mobile application is showing the user status, reports, suggestions for improving the health condition and more. There have been some technologies used for developing SpiMo prototype: Java, MySQL, C++, Bluetooth Low Energy.

Besides the sensors and Bluetooth component, to the Arduino there are also connected some other components: a battery for energy, a piezo speaker for sound notification and a vibration motor.

There have been used eleven pressure sensors in the following manner: 9 in the pillow (placed in a squared matrix of 3) and 2 for the back. The sensors in the pillow are calculating the weight distribution of the person, while the other 2 from the back are detecting if the user is sitting with the back right. Depending on all sensors, we detect if the user is sitting right and we send via Bluetooth the data to the mobile application that is further storing the values in a database and sends notifications. In the mobile application, there is a Boolean variables that is changing when the posture is changing too and depending on it, the application is sending notifications.

### Validation

A posture experiment has been conducted to characterize the performance of the system. At the moment, the prototype consist in the smart pillow (that has an Arduino, sensors, battery and Bluetooth module) and a mobile application which is sending you push notification

regarding the sitting posture, shows you reports and is giving you training exercises.

The experiment was conducted with 8 subjects (4 males and 4 females, age 20-50) and for each we have calibrated the pillow and measured different sitting postures. The testers were alerted each time they have changed the postures from good to bad ones or otherwise. As it can be seen in Figure 2, one that tests the smart pillow has a good posture and receives a pop up on the mobile application with the notification. From the one that participated in the testing phase, 65% of them said that SpiMo is a key for a healthier lifestyle, while 35% were not delighted with the idea, because they considered that are too many notifications and in time they can get stressed because of them.



**Figure 2, SpiMo prototype**

## CONCLUSIONS

Comparing with some other devices that have the same purpose, the good posture of the user, like Lumo Lift for example (a device that you stick it on your body in order to track your activities and posture), SpiMo is a product for people who are working in offices, are spending too much time on a chair, children from schools that are spending time on chairs in classes (they still are in the growing process, so a bad posture might affect their spine later) or simply at home.

SpiMo smart pillow can be used by anyone who would like to take care of their spine and back, provides a pleasant solution, cheap, handy and non-invasive, based on recent technologies, with a great impact on enhancing the quality of life and increasing user's microlives (a unit of risk representing half an hour change of life expectancy) gets improvements, while the micromorts (mortality risk) are decreasing.

Smart Pillow takes care of your, of your health, of your mind and helps you be more productive. It is portable, can be used on any chair.

While people's health is getting improved, money for treatments will be less, Romanian's GDP is increasing, not only because of the less money used on health treatments, but also for increasing the productivity in the IT sector.

The smart pillow is able to teach people how to keep in shape, how to be healthy. Learning how to prevent diseases is the key to a healthy longer life.

As further improvements, the research period will have an important impact of the product, as long as we will aim to develop a helpful product but at lower cost, in order to be affordable for everyone. Another important aspect would be to improve the lifetime of the battery by optimizing the processes in order to not overload the processor and use a wireless technology that is not consuming much.

Machine learning is also a step forward in the development of SpiMo, learning if a keyword that should not be missed in a smart pillow, by making the system learn which are the good posture, which are the poor ones and learning to adapt by itself depending on person which is sitting on the pillow.

## Acknowledgements

The work has been funded by the TRAVEE grant of the Romanian Executive Agency for Higher Education, Research, Development and Innovation Funding - UEFISCDI, Joint Applied Research Projects programme, project number PN-II-PT-PCCA-2013-4-1580.

## REFERENCES

1. Bilal El-Sayed, Noura Farra, Nadine Moacdieh, Hazem Hajj, "A Novel Mobile Wireless Sensing System for Realtime Monitoring of Posture and Spine Stress"
2. J. Van Dieen, M. De Looze, and V. Hermans, Effects of dynamic office chairs on trunk kinematics, trunk extensor EMG and spinal shrinkage, *Ergonomics*, 44, 739-750, (2001), 10.1080/00140130120297. M. Huang, K.
3. Hajizadeh, I. Gibson, and T. Lee, Analysis of compressive load on intervertebral joint in standing and sitting postures, *Technology and Health Care*, 24, 215-223, (2016), 10.3233/THC-151100.
4. W. Y. Wong, M. S. Wong, "Smart garment for trunk posture monitoring: A preliminary study," *BioMed Central*, vol. 3, May 2008
5. Lumo Lift - See <http://www.lumobodytech.com/lumo-lift/>
6. Pomodoro technique - See <https://cirillocompany.de/pages/pomodoro-technique>
7. Wenyao Xu, Zhinan Li, Ming-Chun Huang, Navid Amini, Majid Sarrafzadeh, "eCushion: An eTextile Device for Sitting Posture Monitoring", 2011 International Conference on Body Sensor Networks
8. Steven Haveman, Gijs Kant, "Smart monitoring of worker posture in an office environment"
9. Bilal El-Sayed, Noura Farra, Nadine Moacdieh, Hazem Hajj, Rachid Haidar, Ziad Hajj, "A Novel Mobile Wireless Sensing System for Real-time Monitoring of Posture and Spine Stress"
10. Marko Maslakovic, "Wearables that monitor your posture"
11. Bernd Tessorndorf, Bert Arnrich, Johannes Schumm, Cornelia Setz, Gerhard Troster, "Unsupervised Monitoring of Sitting Behavior"
12. Mengjie Huang, Ian Gibson, Rui Yang, "Smart Chair for Monitoring of Sitting Behavior"

# Transformation of the system sequence diagram to an interface navigation diagram

**William Germain DIMBISOA**  
PhD Student

Laboratory of Computer  
Science and Mathematics  
Applied to Development  
(LIMAD),  
University of Fianarantsoa,  
Fianarantsoa, Madagascar  
dwilliamger@gmail.com

**Thomas MAHATODY**  
PhD

Laboratory of Computer  
Science and Mathematics  
Applied to Development  
(LIMAD)  
University of Fianarantsoa,  
Fianarantsoa, Madagascar  
hasinathomas@hotmail.com

**Josvah Paul**  
**RAZAFIMANDIMBY**  
Full Professor

Laboratory of Computer  
Science and Mathematics  
Applied to Development  
(LIMAD),  
University of Fianarantsoa,  
Fianarantsoa, Madagascar  
jp\_josvah@yahoo.com

## ABSTRACT

The automatic generation of the computer systems such as ergonomic Human Computer Interactions (HCI) always fascinates a computer project designers. For example, the MACAO (Méthode d'Analyse et de Conception d'Applications Orientées-objets) method, which is a participatory method, allows the development of complex object-oriented applications in accordance with the fundamental principles of software engineering. It is based on four models: organizational, structural, behavioral and HCI models. Within HCI models, MACAO offers two types of models: the first model is the Interactive Navigation Diagram (IND) which allows to design and to model the logic of sequence of the interactions between the user and the software; the second model is the Logical Model of HCI which is a translation of the Interactive Navigation Diagram in a particular technology (Windows, WEB, Multimodal ...). In this paper, we perform automatic generation of the Interactive Navigation Diagram from the System Sequence Diagram noted SSD. We use an architecture conforming to the Model Driven Architecture approach (MDA). This study aims at the automatic creation of new models from already existing models. We adopted the Model Driven Engineering (MDE) Approach. Thus, transformation rules are created to automatically generate models.

## Author Keywords

Human Computer Interaction, Interface Navigation Diagram, Model Driven Architecture, Model Driven Engineering, transformation, System Sequence Diagram.

## INTRODUCTION

The objectives of MACAO method are firstly help in the discovery of all the information necessary for the design and development of a software, secondly design the overall and detailed structure of the software in terms of object classes.

This method also define the HCI which is best suited to the needs of users. MACAO Optimize too the maintenance of the delivered software.

It allows to develop reliable, scalable and time-based object-oriented programs. MACAO relies on four-step approach such as: comprehensive analysis, overall design, development and finalization. The four types of models adapted to each area concerned by the software: organizational, structural, behavioral and HCI models. A typical documentation allowing to describe the software in different aspects: technical, maintenance, user and contractual.

In the second step, the method advocates designing the architecture of the software on several levels: functions, business classes, HCI general structure, cutting into prototypes. During HCI general structure design, MACAO uses Interactive Navigation Diagram model.

It is possible to automatically generate this Interactive Navigation Diagram [5] from the sequence diagram. This will facilitate its tasks and minimize the time and load of designing and developing a software. This paper aims to help computer scientist minimize the cost of software development. This article is able to simplify the task of designer while respecting the delay of the project realization.

We will use the MDE approach for the automatic generation of the sequence diagram to the Interactive Navigation Diagram by creating transformation rules. The sequence diagram is the best diagram to highlight the interaction between user and the system. This study is structured as follows: in the first we will talk about the MACAO approach, the MDE approach [6] and its tools. Then we will automatically transform and create the Interactive Navigation Diagram from the sequence diagram. Finally, we will analyze the transformation model from the sequence model to the Interactive Navigation Diagram model. We will test our tool.

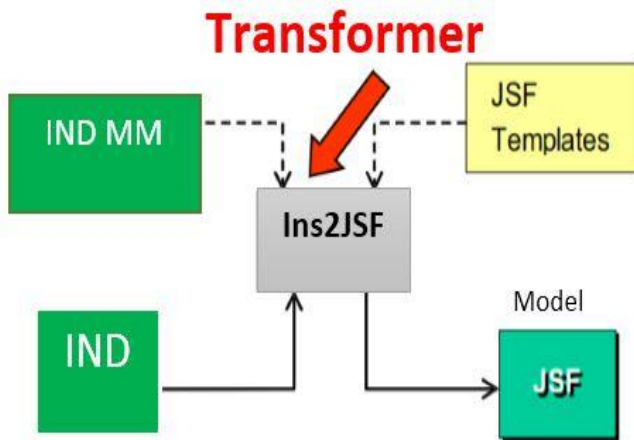


## MACAO AND MDE

In this section, we briefly present the MACAO method and the MDE approach.

### Approach by MACAO

MACAO approach [12] is equipped with a graphic editor of the Interactive Navigation Diagram called VisualSNI who is allowing to manipulate these models. MACAO also allows to transform this Interactive Navigation Diagram model in order to generate semi-automatically a specific model of the HCI. Nicolas Ferry has presented in his research work [12] about the transformation of Interactive Navigation Diagram into Java Server Faces (JSF). The Figure 1 shows this transformation.

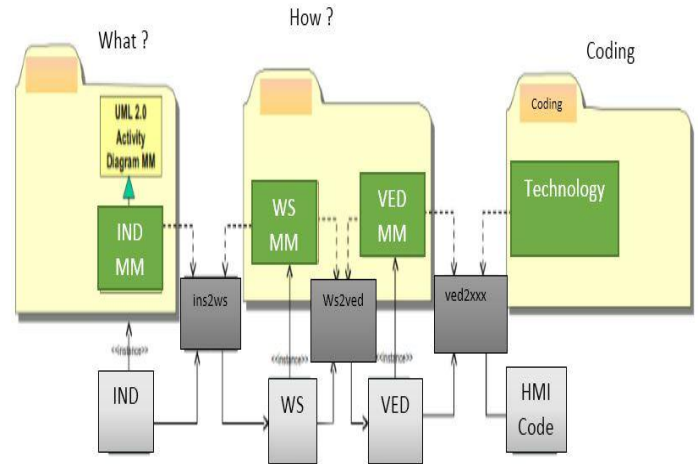


*MM: Meta model*

*IND: Interactive Navigation Diagram*

**Figure 1. A template-based transformer requires the IND metamodel (IND-MM), and a set of templates of transformation in the JSF target technology. By passing a source IND model, we get a mockup with the HCI generated in the form of web pages [12] (modified).**

MACAO has informed a participatory approach from the beginning of the development to the finalization of the software and its maintenance. This method has taken over widely in the design phase the Unified Modeling Language (UML) [7] notation and its diagrams. It has participated in the modeling and the mechanism of the transformation of the models into the MDE [16]. MACAO takes into account the risk factor throughout the software manufacturing process and proposes a criterion for evaluating workloads for the software development [12]. The Figure 2 shows the architecture of the HCI axis for the MACAO approach presented by Nicolas Ferry.



**Figure 2. Overview Architecture of the HCI axis with transformers for the passage of diagrams of different levels. Here MM means Metamodel. This diagram shows the design of a HMI from different levels models of abstraction [12] (modified).**

*IND: Interaction Navigation Diagram*

*WS: Windows Scheme*

*VED: Visual Element Definition*

*MM: Meta model*

In this Figure 2, to answer the WHAT question, the designer uses the Interactive Navigation Diagram model.

The Interactive Navigation Diagram makes it possible to design, model and represent the sequence of a dialogue between the software and the user taking into account the supposed behavior of the user. It is a conceptual model with a high level of abstraction to represent the user's needs and user's functional requirements. The user will be able to navigate between the information offered to him in a seemingly free way, but being constrained to a logic imposed by the designer of the application [12]. The Interactive Navigation Diagram has two types of dialogue unit:

- Basic or elementary dialogue unit;
- Composed dialogue unit by juxtaposition

The Figure 3 shows the structure of the Interactive Navigation Diagram model, the elementary dialogue unit identified by its name.

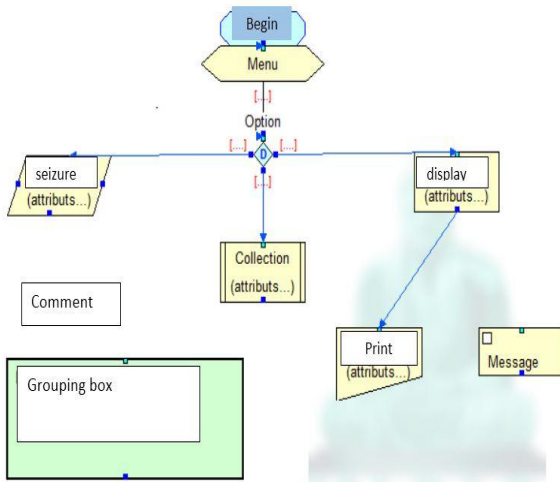


Figure 3. Structure of the IND model, dialogue unit.

**MDE approach**

The Model Driven Engineering approach [6] is the field of computer science that provides tools, concepts and languages to create and automatically transform a models. The architecture directed by the Model Driven Architecture models [10] is a software development approach proposed and supported by the Object Management Group (OMG) which is normalize all of approach. This approach is a particular variant of the MDE. It is structured in three levels. The transition from one model [15] to an another model is ideally carried out automatically. MDA [9] is an approach to implementing an application.

It is also proposed and supported by the OMG. The basic principle of the MDA approach [9, 5] is the development of platform-independent requirement a models and the transformation of these first models into platform dependent models. The first level corresponds to the conceptual model was called the Computational Independent Model (CIM). This model is completely independent of the techniques and the platforms.

It captures the needs and the requirements of users. The second level corresponds to the logical and technical model who is called Platform Independent Model (PIM). It is platform independent and obtained from transforming the CIM level model using transformation rules. The third level is called Platform Specific Model (PSM). It is explicitly the result of the transformation of the PIM model. It represents the final application adapted to a specific platform.

**GETTING INTERACTIVE NAVIGATION DIAGRAM (IND) FROM SEQUENCE SYSTEM DIAGRAM (SSD)**

The creation of sequence diagram [13] is a common and mandatory activity in the design and the development of the software. We aim to reduce the load of carrying out a computer project. The main objective of this article is to automatically transform the sequence diagram to IND.

The approach [12] who is indicated by Nicolas Ferry therefore requires additional tasks for a computer designers and a developers in particular because of:

- semi-automatic transformation of models: existence of human intervention,
- modeling was begun in the Interaction Navigation Diagram model: the trend of the designer is the orientation towards the use of diagrams of the UML notation to model the user interface,
- notion of correspondence between the Interactive Navigation Diagram model and the UML diagram.

In this approach, the transformation of the Interactive Navigation Diagram into an interface model, that is to say the transformation of the PIM level into PSM, has already been carried out by Nicolas Ferry. So this transformation was not taken into account. And, we start designing and modeling directly through the development of the Interactive Navigation Diagram.

**Synoptic diagram of the transformation of sequence diagram to IND**

The manipulation of the Interactive Navigation Diagram model has required additional skills for computer scientists. This results in the loss of time during the realization of a computer project. We have developed appropriate rules to automatically generate models. To support the transformation of sequence diagram into Interactive Navigation Diagram, ATLAS Transformation Language (ATL) [4, 17] is used with its tools. The synoptic diagram of the approach adopted to conduct the research project is presented in the Figure 4.

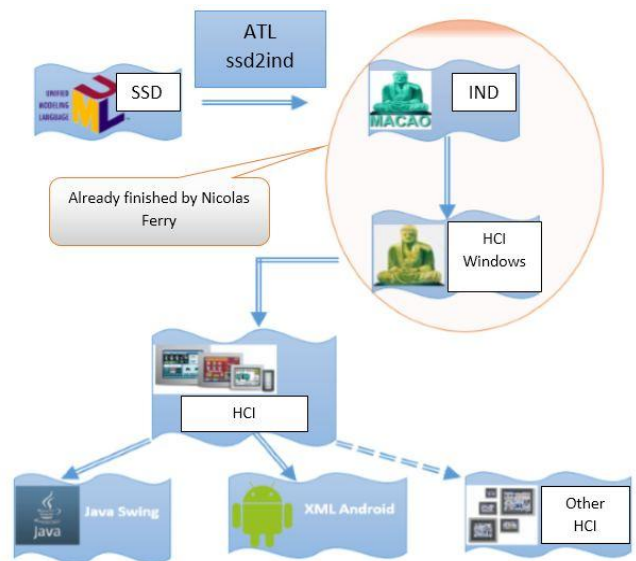


Figure 4. Synoptic diagram of the approach.

In this figure 4, we start the conception by creating SSD and generate the IND. The follow transformation is already finished by Nicolas Ferry to obtain Logical Model of HCI from IND.

ATL is a model transformation language inspired by Query View Transformation. It allows translation between XML Metadata Interchange (XMI) files. It is essential to have an initial and a target metamodels, as well as a mapping file. In this synoptic diagram of the approach, the modeling of an application was begun with the development of the sequence diagram which will be automatically generated in Interactive Navigation Diagram. This Interactive Navigation Diagram will be transformed into an HCI model, and this transformation is already carried out by Nicolas Ferry. In order to minimize the designer's workload, we have created the *ssd2ind.atl* tool to automatically generate the requirements model in logical models of the user interface model.

### Transformation with ATL based on “ecore”

The transition from one model level to an another must follow the transformation rule. These models must be expressed in a modeling language who is defined by a metamodel. A transformation takes as input one or more models of a left part called Left-Hand Side. And it outputs one or more models of a right part called Right-Hand Side. The Left-Hand Side expresses access to the source models, and the Right-Hand Side indicates the expansions so the creation, updating of the target models. Each transformation takes one or more input models and produces one or more output models until obtainable an executable artefacts [6]. An overall view of the model transformation is presented in the Figure 5.

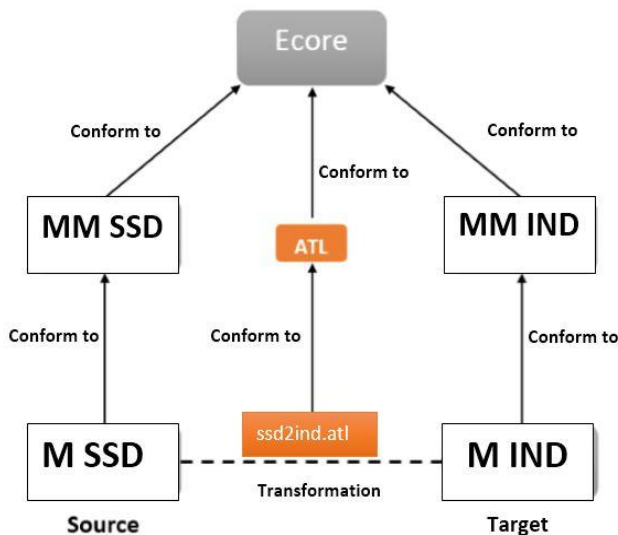


Figure 5. Overview of ATL transformation of sequence system diagram to IND.

M: Models, MM: Metamodels, MMM: Metametamodels

ATL is part of the Eclipse Model-to-Model project. It consists of a transformation language, a compiler with a virtual machine and a tool based on Eclipse.

In the Figure 2, we have the source model conforming to a source metamodel. Modeling is based on the metamodel "ecore". The development of the sequence diagram metamodel begins the transformation. Compliant with this first metamodel, we will create the sequence diagram model as a source model. We have created transformation rules built into *ssd2ind.atl* to automatically generate the INS model as a target model. To manipulate the "ecore" metamodel, the Eclipse Modeling Framework (EMF) is used to process models. This Framework lets us to store templates as files to ensure the persistence. It allows to process the different types of files: conforming to the XMI standard and also in specific forms (Java code).

### Generation of sequence system diagram to Interactive Navigation Diagram

In this study, we need to use the research work [12] of Nicolas Ferry, but instead of designing the IND model, we will automatically generate it from a sequence diagram model. The process of designing and developing software is generally required the creation of the sequence diagram. This study creates a tool for automatically transforming sequence diagram to IND. It highlights the design of the elements of an interaction. To implement the transformation, we will firstly create the sequence diagram metamodel as a source metamodel. The Figure 6 shows a part of the sequence diagram metamodel.

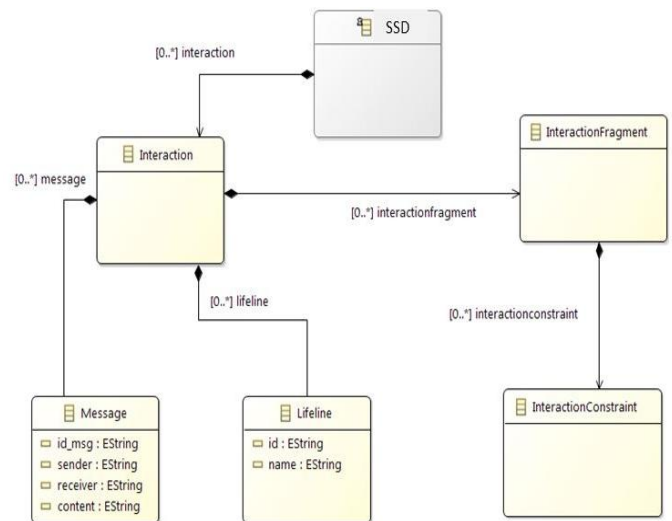
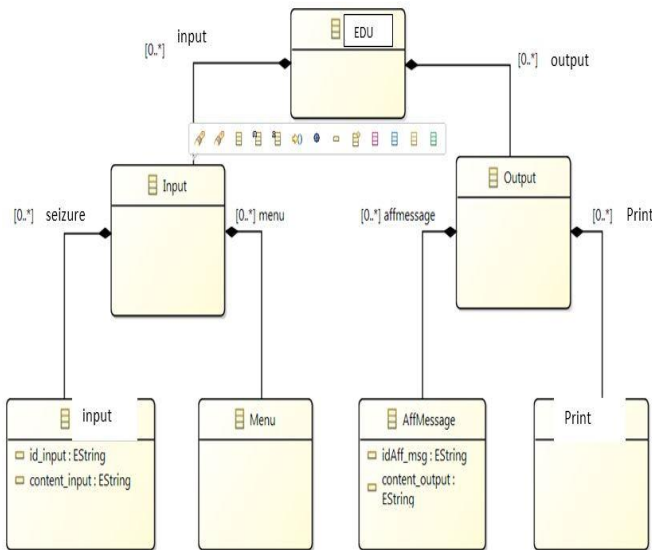


Figure 6. Part of the sequence system diagram metamodel.

This sequence diagram metamodel serves to create a model that is an instance of this metamodel. Secondly, we will create the Interactive Navigation Diagram metamodel as a target metamodel.

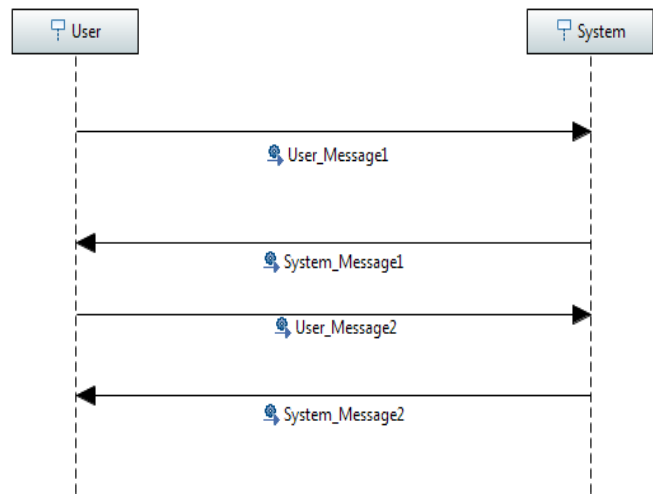
The Interactive Navigation Diagram allows to model and to conceive the logic of sequence of the interactions between the user and the software [6]. It is a conceptual model with a high level of abstraction to represent the needs and functional requirements of users. Interactive Navigation Diagram allows to model the sequence of interaction between the software its user.

Its main objectives are to model the user interface in terms of navigation between the various basic elements, access rights and functional coverage related to the HCI [12]. The Figure 7 shows a part of the Interactive Navigation Diagram metamodel.



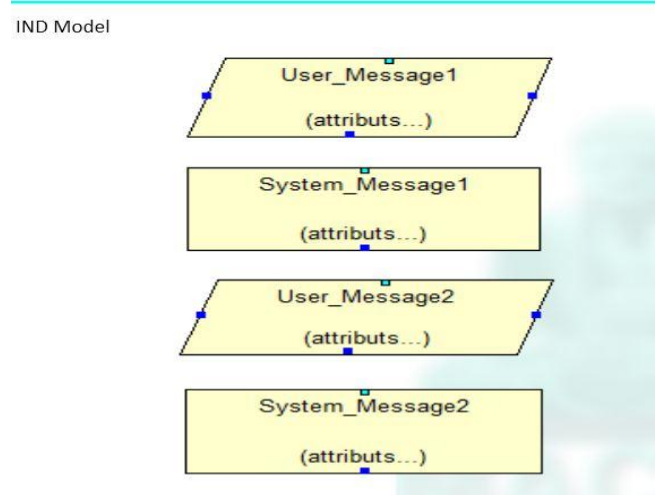
**Figure 7. Part of the Interactive Navigation Diagram metamodel.**

This IND metamodel serves to generate a model which is thus an instance of this metamodel. We will create in the end the sequence diagram model which is an instance of the metamodel of itself. And this model will be automatically transformed into an IND model. Appropriate rules are created to support transformation. The Figure 8 shows the source model called the System Sequence Diagram.



**Figure 8. Model of the System Sequence Diagram.**

The Figure 9 shows the target model called the Interaction Navigation Diagram.



**Figure 9. Model of the Interaction Navigation Diagram.**

From the point of view of the HCI, an interaction takes place either as an *entry in the system* or by *sending to the user*. Dialog box categories are represented with their properties: Display, Collection display, Input, Menu ... [12]. An interaction is mainly represented in two ways:

- Either a message *from the user and directed to the System*: this message is transformed into a dialogue unit called *Input* in IND;
- A message *from the system to the user*: this message is transformed into a dialogue unit called *Output* or *AffMessage* in IND.

The representation in the Figure 8 and the Figure 9 implements the correspondence between the interaction message of the sequence diagram model and the dialogue unit of the IND model.

## DISCUSSION

The purpose of modeling an interactive application is to allow software developers to define the models that they use for an application. We obtain the Interaction Navigation Diagram model as a result of this study. This model is automatically generated from the sequence system diagram to minimize loads and facilitate the tasks of interactive application designers. The purpose of this article is to save time by automatically creating the Interaction Navigation Diagram from the sequence system diagram. This sequence system diagram is not enough to design an user interface, but to highlight the interaction between the user and the system it is interacting with, it is better to first build the sequence diagram.

The modeling of HCI according to the architecture proposed by Nicolas Ferry was begun directly with the development of the Interaction Navigation Diagram. This study encourages the start of design by creating the sequence system diagram to capture the needs of users. On the one hand, the modeling of HCI can be started by directly developing the Interaction Navigation Diagram model using the VisualSNI tool. But, the direct creation of the Interaction Navigation Diagram was required additional tasks for the computer scientists. Then, we solicit to automatically generate the Interaction Navigation Diagram model from the sequence diagram model. The aim of this study is to encourage the designer and the software developer to realize evolving programs on time. We are used the UML [10] notation to model the development. The Interaction Navigation Diagram model have many advantageous interaction elements for HCI.

Nicolas Ferry has presented only a tool for transforming the CIM level to the PIM level, in this case the transformation of Interaction Navigation Diagram into an interface model. For the transformation of the sequence system diagram to Interaction Navigation Diagram, the elements of the Interaction Navigation Diagram are no longer specified to the interactions validated from the keyboard or mouse. For the automatic generation of the Interaction Navigation Diagram from the sequence system diagram by the ATL tool, we used the XMI format of the Papyrus, VisualSNI. There are a lot of Framework but we adopt EMF [2] for usefull especially in model transformation in Eclipse development environment.

There are a lot of similar approach such as approach [1] by Arnaud Brossard based on PERsonalization and COncceptual MOdeling Method (PERCOMOM) [3] and approach based on MACAO.

The table 1 shows the comparison between the similar approaches.

Approach	Transformation rules	Tools	Transformation
[1] based on PERCOMOM	Does not having rules	Insufficient tools	Semi-automatic
[12] based on MACAO	Having rules	Insufficient tools	Semi-automatic
Proposed approach	Having rules	Use of created tools	Automatic

**Table 1. Comparison of the similar approaches.**

Thus, this article offers the following opportunities:

- Time saving: our study makes it possible to obtain Interaction Navigation Diagram immediately from sequence system diagram, it also makes it possible to minimize the costs of carrying out a computer project;
- Reusability: Transformation rules are customizable and templates are reusable, the generated model is specific for a platform;

We have tested our tool to realize the transformation.

## CONCLUSION

The specific architecture of this study conforms to the MDA approach. It allows to automatically generate target models from existing conceptual models. The research strategy is based on the design, the modeling and the realization of a Logical Model of HCI. We have defined a HCI more adapted to the needs of the users by using approach which based on MACAO method. The approach makes it possible to simplify as much as possible the modelization of an application. The main objective is to solicit the designer to simplify the process of the realization of computer project. The approach consists in helping computer scientist to facilitate tasks, reduce and minimize the time required to complete an application. This study allows to automatically generate Interaction Navigation Diagram from the sequence diagram to facilitate the tasks of the HCI designers. In this way, it reduces loads and saves the time needed to produce an application. On this study, we will consider the retro engineering in the future work.

## ACKNOWLEDGMENTS

I would like to express my sincere thanks to all those who helped us in designing and carrying out this paper.



## REFERENCES

1. Arnaud, B., Mourad, A. and Christophe, K. Context Awareness and Model Driven Engineering: A Multi-Level Approach for the Development of Interactive Applications in Public Transportation, 2008.
2. Cédric, B. and Alfonso, P. Model Differences in the Eclipse Modelling Framework, 2008.
3. Firas, B. and Mourad, A. Tool support for conceptual models of the PERCOMOM platform applied to transportation domain.
4. Frédéric, J., Freddy, A., Jean, B. and Ivan K. ATL: A model transformation tool, 2017.
5. Igor, S., J., V. Introduction to Model Driven Architecture (MDA), 2007.
6. Jean, B. In Search of a Basic Principle for Model Driven Engineering, 2004.
7. Jean, B., C. and Nicolas, F. Modélisation PIM des IHM : VisualSNI, 2007.
8. Jean, B., C. and Nicolas, F. SNI: High Level Model for HMI Design and Mock-up, 2008.
9. Jean, B., From Object Composition to Model Transformation with the MDA, 2001.
10. Kuznetsov, M. B. UML Model Transformation and Its Application to MDA Technology, 2007.
11. Li, D. and Li, D. An Approach to Formalize UML Sequence Diagrams in CSP, 2012.
12. Nicolas, F. Formalisation des modèles de la méthode MACAO et réalisation d'un outil de génie logiciel pour la création d'interfaces homme-machine, 2008.
13. Sawprakhon, P. and Limpiyakorn, Y. Sequence Diagram Generation with Model Transformation Technology, 2014.
14. Tisi, M., Martinez, S. and Choura, H. Parallel Execution of ATL Transformation Rules, 2013.
15. Tom, M., A Taxonomy of Model Transformation, 2006.
16. Wided, B., Adel, M., Lassaad, B., Sirine, R. and Mourad, A. An MDE Approach for User Interface Adaptation to the Context of Use. HCSE 2010, LNCS 6409, pp. 62–78, 2010.
17. Yassine, R., Youssef, H. and Abdelaziz, M. Model Transformation with ATL into MDA from CIM to PIM Structured through MVC, 2016.



# Mapping Questions to Ontology Components for Question Answering over Linked Data

Marius Valeriu Stanciu, Eugen Vasilescu, Stefan Ruseti, Traian Rebedea  
University Politehnica of Bucharest, Faculty of Automatic Control and Computers

Bucharest, Romania

{marius.stanciu1710, eugen.vasilescu}@cti.pub.ro, {stefan.ruseti, traian.rebedea}@cs.pub.ro

## ABSTRACT

This paper describes a statistical method to identify ontology components within natural language questions. The main purpose for this step is to improve question answering systems over linked data by reducing the ambiguity in the subsequent matching and query generation steps. To accomplish this task, we have trained a Conditional Random Field (CRF) classifier to label sentence tokens with the core data elements of the DBpedia ontology. The classifier was trained on a manually annotated corpus labelled with ontology elements for each token. Several features were investigated for the classifier and the results (F1=0.92) prove that this task can be successfully solved using the CRF tagger.

## Author Keywords

Natural Language Processing; Question Answering; CRF Tagger; Ontology; DBpedia

## ACM Classification Keywords

I.2.7 Artificial Intelligence: Natural Language Processing

## INTRODUCTION

Knowledge bases are playing an increasingly important role in enhancing the intelligence of the Web. Such an example is DBpedia [1], which collects information from Wikipedia and organizes it in a machine-friendly manner. The problem of question answering over linked data is not a trivial task, even with all the available information. One of its hardest sub-tasks is matching tokens with their corresponding elements in the ontology. Having this done, processing the question intent and translating it into SPARQL queries will not be so far from reachable anymore. DBpedia organizes its data elements into 3 core categories:

- **Entities:** the subjects of description within the ontology. Each one has a dedicated web page, and corresponds to a Wikipedia page from which the information has been aggregated. Descriptive data (literal dates, strings, numbers) or other entities are linked through properties.
- **Types:** classes of entities structured hierarchically. The ontology by itself contains a limited number of types consisting of relatively vague nouns (e.g. *Person*, *Religion*, *City*) but types from multiple other sources have been added. For example, Yago [11] types usually represent more specific item categories such as *President* or *American Lawyers*, but they are considerably less regulated and consistent.

- **Properties:** they are usually predicate-like structures that link additional data to entities. Their representation is the most unpredictable as they do not follow any obvious syntactic or semantic patterns.

## RELATED WORD

### Named Entity Recognition

One popular example of a well-known and similar task to ours is called Named Entity Recognition (NER). It determines and classifies named entities from a text into predefined categories like persons, organizations, date, time, or money. For example, Stanford NER [5] uses the Maximum Entropy Markov Model algorithm [8]. Our task is a bit more complex because the difference between properties and types is not so obvious, but the approach to solve the problem should not be too different

### Question Answering over Linked Data

Like we mentioned in the introduction, the ultimate goal of this model is integration within a question answering system, thus it is worth mentioning a few examples and what impact our algorithm could have.

One example is QAnswer [10], which uses a pipeline architecture for processing the questions. The first and most crucial step in the algorithm is trying to detect the elements within the question (they used individuals, types and properties) and map them to their corresponding resource in the ontology. After this step is completed and relationships between elements are made, a SPARQL query is built and, using a Virtuoso-opensource endpoint (<https://github.com/openlink/virtuoso-opensource>), their model can generate the answer. They have developed three separate mapping algorithms based on the characteristics of every ontology element and then picked the most appropriate sequence. Since they had no information about what the words could represent, they had to construct different interpretations of the questions and, based on some scoring algorithm, they picked the best one. This approach didn't produce the desired result all the time so, some questions were compromised. If they had additional information about the question (like our algorithm could provide), they would have probably had better results.

Another question answering tool worth mentioning is Xser [12], which was ranked first in QALD-4 and QALD-5 (see <https://qald.sebastianwalter.org/>). They have a two-layered

architecture, where the first one tries to label all phrases with one of the following tags: entity, relation, type and variable. This is the part that interests us the most because, for doing this, they also had to train a classifier, more precisely a structured perceptron [3]. The algorithm shows good results but we believe that it could benefit from the use of a sequence labeling algorithm. Moreover, using additional features for the classifier could further improve the precision. The second step of their algorithm is to map the discovered elements to a knowledge base, like DBpedia and construct the query.

### PROPOSED SOLUTION

The main challenge consists in finding correlations between the 3 core data elements described in the introduction and natural language patterns in questions.

#### Labels

Our algorithm should be able to classify every word as part of an ontology element (entity, type, property) or mark them as irrelevant. We needed to find a standard for doing this and, since the elements could contain more than one word (e.g. *Barack Obama*), we chose the IOB tagging standard (Inside, Outside, Beginning). We defined the following tags: EB (entity beginning), EI (entity inside), TB (type beginning), TI (type inside), PB (property beginning), PI (property inside) and N (none).

#### Examples

To illustrate the labeling task, let us consider the following examples:

- *Which is the largest city in Australia?* In this example, we have one entity: *Australia*, one type: *city* and one property: *largest*. The other words cannot be mapped to an element in the ontology so we can consider them as irrelevant. So, the tagging sequence for the above example becomes (N, N, N, PB, TB, N, EB).
- *What did Bruce Carver die from?* In this question it is very clear that we should consider *Bruce Carver* as an entity, but it may be difficult to choose the property. Our goal for this algorithm is to identify ontology elements in phrases so that instantiating them afterwards is easier. In other words, the labels should ease the mapping of *die from* to *dbo:deathCause*. The problem here is that *Bruce Carver* has multiple other similar properties like *dbo:deathDate*, *dbo:deathPlace* or *dbo:deathYear*. By choosing *die* as the only word for the property, distinguishing between the four options might prove impossible. As a result, when prepositions offer additional contextual information, they should also be included. Concluding, the correct tagging sequence becomes (N, N, EB, EI, PB, PI).
- *Give me all video games published by Mean Hamster Software.* This is a more complex example, containing all possible tags. Firstly, we can identify *Mean Haster*

*Software* as an entity. It contains 3 words, so the tags should be EB, EI, EI. Additionally, we have one composite type: *video game* and one property: *published by*, constructed using the same rules as the above example. Then, the correct tagging sequence becomes: (N, N, N, TB, TI, PB, PI, EB, EI, EI).

Our corpus consists of 600 questions which were manually annotated with the previously defined tags. We have used Gate [4] for simplifying the annotation process, while the questions were taken from multiple sources: QALD-6, QALD-7 and WebQuestions Semantic Parses Dataset (<https://www.microsoft.com/en-us/download/details.aspx?id=52763>). From this corpus, 400 questions were used for the training process, while the rest were kept for validation.

#### Choice of Model

We considered several algorithms used to solve labeling problems in Natural Language Processing (NLP), including Hidden Markov Models, Maximum Entropy Markov Models, Conditional Random Fields, and Neural Networks.

The Hidden Markov Model (HMM) [9] is a well-known sequence labeling algorithm, but it's not the most reliable one because it has direct dependencies only between states and their direct observations. An improvement to this algorithm brings the Maximum Entropy Markov Model [8], which is inspired from the Hidden Markov Model and the Maximum Entropy theory. It models dependence between each state and the full observation sequence explicitly, but it suffers from the label bias problem (states with low entropy transition distributions tend to ignore their observations).

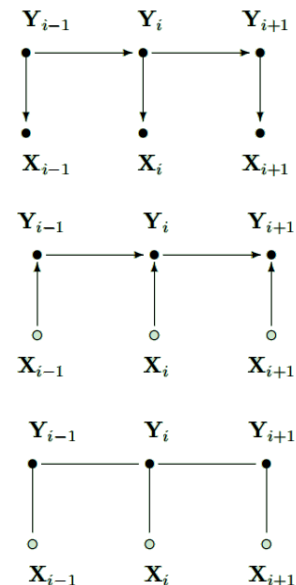


Figure 1: Difference between HMM (first), MEMM (second) and CRF (last) graphical models. From Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data [7]

Conditional Random Fields (CRF) [7] were designed to overcome the label bias problem while also taking into consideration the full context for the predictions. The differences between those three models can be clearly observed in Figure 1.

The other option that we considered was the Long Short-Term Memory (LSTM) neural network [6]. Its obvious advantage is that it automatically extracts features and it has good results in many areas. What made this approach impracticable for us is the big corpus required, one that we don't have. So, considering the arguments presented above, we chose the Conditional Random Field classifier, a choice which was proven to be a good one.

### Choice of Features

Various features have been tested both individually and in relation with others. The following section is a description of the tested features and their influence over the results.

Considering entities are usually proper nouns (e.g. names of personalities, cities or organizations) within the ontology, identifying them is obviously similar with the Named Entity Recognition (NER) problem. Making use of thoroughly trained and tested NER models is definitely going to help at this stage. For this purpose, we have tested two popular models:

- Stanford 3-class NER tagger [5]: trained on various data sets, with generic classes: *Location*, *Person* and *Organization*. Even though there are 4 and 7-class versions of this labeling model, for the purpose of simply identifying named entities, the further classification would only add unnecessary complexity. Using this feature by itself results 82% accuracy over entities. However, as suspected, it does not contribute to identifying types and properties very much (less than 10% accuracy).
- Spacy NER tagger (see <https://spacy.io/docs>): a more modern model that shows minor improvements (about 5%) over the previous one.

On the other hand, types and properties do not follow such predictable patterns, making their identification less trivial. Analyzing DBpedia types reveals that they are usually vague common nouns (e.g. *person*, *football player*, *athlete*) and they are in close relation with entities or *WH* question words (e.g. *Where*, *What*, *Who*). Additionally, properties are usually predicate-like structures, connecting a subject to an object. Both types and properties are highly flexible regarding formulation and they are prone to polysemy, making string matching techniques less effective. This implies that part of speech and syntactic dependency information is essential to their identification. Again, this knowledge is available as widely-popular models in form of Part-of-Speech (POS) tagging and dependency parsing algorithms. The most promising models we considered are:

- Stanford POS tagger part of the NLTK package [2]: individual tests show an average of 43% labeling precision.
- Spacy POS tagger: this tagger offers two format variants, a coarse-grained one, based on the 12-class Google Universal POS Tags, and a fine-grained one, based on the 36-class Penn Treebank specification. Considering the relatively small train set, the lower detail specification shows minor improvements over the fine-grained one, averaging at 49% accuracy. However, upon future train set expansion, the more detailed one should be considered.
- Spacy Dependency Parser: used to extract the phrase type and the graph direction (whether tokens are the governing or the dependant part of a relation). This feature by itself shows modest results (average of 23%) However in relation to the others, it adds a considerable contribution.

Additionally, adding the following string based features showed significant improvements over the average precision.

- Lemmatized tokens: reducing words to their lemmatized version implied the emergence of more common text patterns, essential indicator for property and type recognition.
- Token index within sentence divided by the length of the sequence: types have the tendency to appear at the beginning of questions (excepting syntactic inversions, which are quite common in interrogative sentences) while properties and entities exhibit less regular positions.
- Word suffixes: similar words tend to have the same suffixes. For example, *son* is a suffix that appears mainly in person names (e.g. *Johnson*, *Anderson*, *Jackson*, *Dawson*), while *ies* commonly appears in tokens defining types (e.g. *cities*, *movies* or *countries*)
- Capital first letter: every word that starts with an uppercase letter has a great chance to be an entity. This helps the algorithm to be more confident about entity predictions.

Considering the highly context-dependent nature of the discussed elements, the relationship between the above-mentioned indicators is preserved and enforced using various templates, illustrated in formulas (1), (2), and (3).

- Unigrams:

$$w_{i-1}, w_i, w_{i+1} \quad (1)$$

- Bigrams:

$$(w_i, w_{i+1}), (w_{i-1}, w_i) \quad (2)$$

- Trigrams:

$$(w_{i-2}, w_{i-1}, w_i), (w_{i-1}, w_i, w_{i+1}), (w_i, w_{i+1}, w_{i+2}) \quad (3)$$



Tag	Precision	Recall	F1-Score	Support
EB	0.97	0.90	0.93	125
EI	0.90	0.86	0.88	73
N	0.96	0.97	0.97	369
PB	0.89	0.91	0.90	99
PI	0.83	0.89	0.85	53
TB	0.82	0.85	0.83	84
TI	0.81	0.73	0.75	13

**Table 1. Precision and recall for each labeled tag from the test set. Average overall F1-score is 0.9**

## RESULTS

After a thorough analysis of feature combinations and templates, we finally settled on using mainly n-grams of Spacy POS, NER and dependency labels together with the above-mentioned string-based features. The model results are illustrated in Table 1. As can be observed, the highest confidence is achieved for entities with over 90% accuracy, result which was expected due to their predictable structure (almost every entity starts with a capital letter, for example). Types and properties can be easily confused, but the accuracy doesn't drop below 80%. Features like the POS tags and the dependency type returned by the dependency parser were essential in obtaining these results.

## CONCLUSIONS

In this paper, we have proposed to use a Conditional Random Field (CRF) model to label question tokens as core data elements from the DBpedia ontology in order to improve question answering systems over linked data, such as DBpedia. The features we used are mostly generated by widely-popular Natural Language Processing tools such as POS taggers, NER taggers and dependency parsers. The ultimate goal of this method is to remove ambiguity and improve further matching in the context of question answering over linked data. Even though training and testing were done over a relatively small manually annotated corpus, the model shows good results with average labeling accuracy of 92%.

## ACKNOWLEDGMENTS

This research was partially supported by University Politehnica of Bucharest through the Excellence Research Grants Program UPB-GEX 13/30.09.2016.

## REFERENCES

1. Soren Auer, Christian Bizer, Georgi Kobilarov, Jens Lehmann, and Zachary Ives. 2007. DBpedia: A Nucleus for a Web of Open Data. In *In 6th Int'l Semantic Web Conference, Busan, Korea*. Springer, 11–15.
2. Steven Bird. 2006. NLTK: the natural language toolkit. In *Proceedings of the COLING/ACL on Interactive presentation sessions*. Association for Computational Linguistics, 69–72.
3. Michael Collins. 2002. Discriminative training methods for hidden markov models: Theory and experiments with perceptron algorithms. In *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10*. Association for Computational Linguistics, 1–8.
4. Hamish Cunningham, Diana Maynard, Kalina Bontcheva, Valentin Tablan, Niraj Aswani, Ian Roberts, Genevieve Gorrell, Adam Funk, Angus Roberts, Danica Damljanovic, Thomas Heitz, Mark A. Greenwood, Horacio Saggion, Johann Petrak, Yaoyong Li, and Wim Peters. 2011. *Text Processing with GATE (Version 6)*. <http://tinyurl.com/gatebook>
5. Jenny Rose Finkel, Trond Grenager, and Christopher Manning. 2005. Incorporating non-local information into information extraction systems by gibbs sampling. In *In ACL*. 363–370.
6. Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation* 9, 8 (1997), 1735–1780.
7. John Lafferty, Andrew McCallum, Fernando Pereira, and others. 2001. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *Proceedings of the eighteenth international conference on machine learning, ICML*, Vol. 1. 282–289.
8. Andrew McCallum and Dayne Freitag. 2000. Maximum entropy markov models for information extraction and segmentation. Morgan Kaufmann, 591–598.
9. Lawrence R Rabiner. 1989. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE* 77, 2 (1989), 257–286.
10. Stefan Ruseti, Alexandru Mirea, Traian Rebedea, and Stefan Trausan-Matu. 2015. QAnswer-Enhanced Entity Matching for Question Answering over Linked Data.. In *CLEF (Working Notes)*
11. Fabian M Suchanek, Gjergji Kasneci, and Gerhard Weikum. 2007. Yago: a core of semantic knowledge. In *Proceedings of the 16th international conference on World Wide Web. ACM*, 697–706.
12. Kun Xu, Sheng Zhang, Yansong Feng, and Dongyan Zhao. 2014. Answering natural language questions via phrasal semantic parsing. In *Natural Language Processing and Chinese Computing. Springer*, 333–344.

# Real-time Guided Procedural Terrain Generation

**Sima Vlad Grigore**

Technical University of Cluj-  
Napoca

Str. G. Barițiu 28, 400027,  
Cluj-Napoca, România  
simavlad\_10@yahoo.com

**Adrian Sabou**

Technical University of Cluj-  
Napoca

Str. G. Barițiu 28, 400027,  
Cluj-Napoca, România  
adrian.sabou@cs.utcluj.ro

## ABSTRACT

This paper presents an application that allows real-time large terrain generation using crude user input regarding the terrain's height features. Our technique uses a preprocessing phase to generate plausible terrain out of crude bitmaps and achieves real-time performances by splitting the large terrain into chunks and processing and displaying only the chunks surrounding the user, while maintaining the height features suggested in the original crude input.

## Author Keywords

Terrain synthesis; Real-time; Heightmap; Worley noise; Perlin noise; Filters.

## ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI):  
Miscellaneous.

## General Terms

Algorithms; Terrain models.

## INTRODUCTION

Creating and modeling realist environments, has become over the last decades one of the most disputed and discussed aspect in computer graphics. Given the fast evolution that technology is facing these days, and the significant hardware progress in graphics, more and more detailed terrains are developed using a procedural approach. Since 1982 when it found its first commercial use in the well-known Star Track series by generating alien planets, the concept of creating content based on algorithms or rules, has faced a rapid and impressive growth, being used in large and diverse areas nowadays.

Procedural content generation (PCG) represents the process of creating content with limited or absent implication from the user. This doesn't always mean that the user won't be able to affect in anyway the result of the generation part, but that this process will mainly rely on the randomness of the algorithms involved. Based on the stage in which it is created, we can distinguish two different types of procedural content generation. The first one is offline PCG, and it's called this way because all the content is generated during the development stage and the result comes with all the features already implemented. This is a method that is very safe regarding faults, and has a lower demand for the algorithm because, in both cases, all the errors or the results

that don't fit a particular expected model can be corrected by hand.

The second type of PCG is the online version, which as the name suggests, involves creating terrain at startup, also referred to as real-time generation of content. Despite the fact that in this scenario the algorithms are far more complex, it can offer various opportunities for generating infinite terrain, creating an unforgettable experience for the users.

A common agreement in the field of PCG is that neither the procedural nor the model base approach fulfill the needs of the user in what concerns a realistic representation of the environment. On one hand, using solely a procedural approach will leave the terrain with uncharacteristic features and, on the other hand, by employing the model-based approach, terrain lacks realism. In order to solved this drawback, a compromise is made between these two possible approaches, leaving many windows of opportunities for creating suitable content.

The same approach is utilized in this paper, relying on both the procedural method and the model based one in order to obtain realistic output that follows the user's input. In the context of our application, using only the user's input would provide a terrain that lacks detail and, as there are no pre-conditions for the level of detail for the initial bitmap, there might be unrealistic transitions between adjacent zones that is not normal in a natural environment. Using the other approach, a purely procedural one, would result in a terrain which might not meet the user's expectations and also might developed uncharacteristic features.

## RELATED WORKS

Olsen [1] created real time terrain, but the paper was not fully focused on performance and how to obtain it, but rather it was focused on presenting ways of using and improving algorithms for erosion. His real-time generation has a loading time, which can be an acceptable way of developing a real-time application, as many games nowadays have an allocated loading time, in which they generate some pre-required data or just fetch them from memory. Our application behaves in a similar way, in order to generate the features points. In what concerns the creation of terrain Olsen used a  $1/n$  noise. Then he combined this with Voronoi diagrams, which is similar to

Worley noise which we used as a base algorithm to develop a suitable one for our needs. Lastly, he used improved thermal erosion, modified to simulate the properties of hydraulic erosion.

Génevaux et al [2] generated content using the hydrological erosion, and using a hydrology-based method in order to represent it. The construction of the algorithm starts by using the input provided by the user. The input is composed of the following elements: terrain outline, river mouths and other parts of the river. The aim of the algorithm is to form a complete river network, kept under the form of a graph, starting from some basic elements and sketch provided by the user. When the algorithm is completed, blocks such as junctions, springs, deltas, and river trajectories are created, and later used to help rendering the final terrain.

Mangra et al [3] present an approach to terrain synthesis from minimal-detail user-provided heightmaps. There is no assumption regarding the level of detail provided, in order to allow users without access to powerful heightmap tools and/or resources to generate useable terrain based on a self-provided crude feature plan. They present the issues stemming from a lack of detail in user input, notably sharp altitude increases and oversimplified feature edges, and proceed to elaborate on using the terrain synthesis algorithm to solve the issues and create a level of detail that more closely resembles realistic terrain models. The algorithm pipeline is presented and parametrized to show how the user can influence the resulting model.

Zhou et al [4] present an example-based system for terrain synthesis. In their approach, patches from a sample terrain (represented by a height field) are used to generate a new terrain. The synthesis is guided by a user-sketched feature map that specifies where terrain features occur in the resulting synthetic terrain. Both the example height field and user's sketch map are analyzed using a technique from the field of geomorphology. The system finds patches from the example data that match the features found in the user's sketch. Patches are joined together using graph cuts and Poisson editing. The order in which patches are placed in the synthesized terrain is determined by breadth-first traversal of a feature tree and this generates improved results over standard raster-scan placement orders. Their technique supports user-controlled terrain synthesis in a wide variety of styles, based upon the visual richness of real-world terrain data.

Schneider et al [5] interactively synthesize artificial terrains using procedural descriptions. They present a new GPU method for real-time editing, synthesis, and rendering of infinite landscapes exhibiting a wide range of geological structures, building upon the concept of projected grids to achieve near-optimal sampling of the landscape. They describe the integration of procedural shaders for multifractals into this approach, and propose intuitive options to edit the shape of the resulting terrain. The method is multi-scale and adaptive in nature, and it has

been extended towards infinite and spherical domains. In combination with geo-typical textures that automatically adapt to the shape being synthesized, a powerful method for the creation and rendering of realistic landscapes is presented.

## **REAL-TIME GUIDED TERRAIN GENERATION**

While purely procedural terrains have a tremendous amount of detail over terrains generated based on models, they still lack the control the user can have over the final result of the process. In order to overcome this issue, this paper proposes a simple but efficient algorithm that would help the user control the major elements from the terrain he wants to create, by being able to offer as input a crude grayscale bitmap where he can mark the main elements from the landscape.

Regarding our application, we can divide it into two main parts that are going to be discussed later on, in this paper. First, we focus on the procedural part that will contain three important steps: edge smoothing, adding detail using a type of noise and then the process of smoothing the noise created in the previous step. The second part of this paper focuses on generating terrain in real time by creating new chunks of terrain as the user moves through the scene. This part will apply the steps described in the first one, and as the user moves, new chunks of terrain will be generated, while the last ones will be deleted from the scene, all of these steps being controlled by a predefined range or radius.

### **Procedural Terrain Generation**

We are going to explain step by step each of the aforementioned steps, starting with edge smoothing. This step is required because, as specified earlier, the user input has no restrictions in what concerns the level of detail it must provide. So, there is quite a chance that the inexperienced user might provide a black image with a white square in the middle. The problem here is that the transition between the plain represented by the black color and the mountains represented by the white color, would be a 100% height transition and it would look extremely unnatural. In order to avoid this, we use Worley Noise [6] as a source of inspiration, an algorithm designed to help smooth the boundaries (edges) of the input bitmap. The idea of the algorithm is to generate uniformly distributed random features points all over the map. In order to ensure this will happen, the initial bitmap was split into equal grids and, in each grid, a particular number of seed points was generated. This way we ensure that they are distributed uniformly on the surface. Then, in order to compute the actual height value of the current pixel, we compute a weighted sum of values from the closest  $N$  features points. Thus, each of the feature points will contribute with a percent directly proportional with the distance between it and the point whose value is being computed.

The second phase is represented by the addition of detail to the equalized bitmap. This step is absolutely necessary

because the user's input might not provide enough details for the map to be relevant and look like a real environment. As presented in the last step, if the user decides to introduce a crude bitmap with only two colors and some random geometrical figure, then the terrain will look unnaturally flat. To overcome this limitation, a type of coherent noise must be used in order to obtain suitable results that also respect the users input. For this task, we picked Perlin Noise [7], a type of coherent noise capable of adding new and significant details. Perlin Noise uses a gradient approach, by computing gradient vectors in each corner of the square, if we refer to the 2D version, or cube, if we refer to the 3D version, then uses linear interpolation between these gradient vectors in order to compute the height value of the current point.

Last, but not least, the step that we are going to discuss is the smoothing we will apply to the height map that was generated so far, by using two different types of digital filters, the mean and median filters [8][9]. This step is extremely important, as Perlin Noise might have made some unwanted changes with a particular point, meaning it could have assigned it a greater value than expected or a lower one. In this case the value of the point will be recomputed based on his neighbors' heights by using these two types of filters. For the mean filter, we use a kernel of size 3x3, computing the mean of all surrounding neighbors' height values as the current height. The other type of digital filter used is the median filter, which sorts neighbors and then picks the value that lays at the center of the interval. After this step is completed the height map is ready to be rendered, and this happens by applying it to multiple terrain game object.

### **Real time guided generation**

The second step of our method is represented by the real-time generation, which involves the creation of new terrain as the user moves though the environment. The algorithm works as follows: it places the user in the scene's origin and it crops 9 chunks from the input provided by the user, transforming them into terrain later. The main idea is to use a tile base approach, by utilizing 3x3 matrix of chunks, placing the user in the middle and updating it as the user moves thought the scene. The initial position in the input provided by the user is in the center of the crude bitmap. In order to achieve this functionality, we use the center of the image as a pointer, that will be later on updated based on users' decisions.

Then, on the chunks that were just taken from the bitmap, we apply the first step of the algorithm in order to create a viable height map that still respects the input that the user provided. At this point we have a height map that represents part of the terrain that the user wanted. Now, as this application offers the possibility for the user to experiment real time generation, if the user decides to move in a random direction on the X and Z axes, the application will intercept this. It uses a function that checks in each frame

the current position of the user and, if the user travels a certain distance from the starting point, the pointer that indicates the user's position in the bitmap will be updated using a predefined value and taking into account the direction in which the user chose to walk. As mentioned in the previous sections, the application is not meant to generate infinite terrain, the size of the result being directly influenced by the size of the bitmap given by the user. When the user reaches the end of the environment, it means that the user has reached the edge of the input he provided so the application won't generate any more terrain and the user will have to either move in another direction or get stuck as a collision detection will be implemented to prevent the user from falling,

## **EXPERIMENTS**

### **Validating results**

Initial testing was done to prove that the algorithm does indeed successfully provide a detailed model of plausible terrain. Our real-time generation algorithm was applied both using a pre-refined heightmap and a crude user created bitmap to guide it.

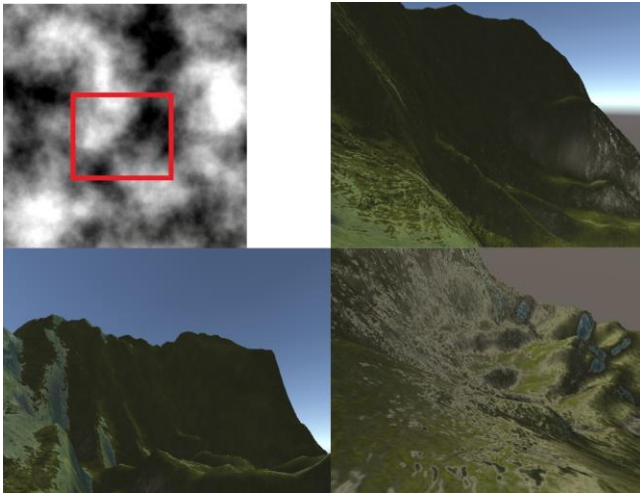
Figure 1 shows the result when applying our algorithm using a pre-refined heightmap as a source. The red square represents the area of terrain that is generated around the user's position and the screen captures illustrate what the user sees when pointing north, south and west. As the user moves around the scene, the generated landscape will change accordingly, but still preserve all the height characteristics in the heightmap.

### *Guided terrain generation*

Since one key aspect of this application is the possibility to guide the algorithm using crude user generated bitmaps, the second test case was built around such input. Figure 2 shows the resulting terrain when applying our technique using a crude bitmap as a source. Again, the red square represents the area of terrain that is generated around the user's position and the screen captures illustrate what the user sees when pointing north, south and west. As the user moves around the scene, the generated landscape will change accordingly, but still preserve all the height characteristics suggested in the original crude bitmap source.

### *Performance considerations*

Performance is directly connected with the size of the image, because the feature points algorithm divides the image into smaller grids of size 4x4 and generates a fixed number of random samples that are later used to compute the height values for each pixel. The experiments made were with a 128x128 size pixel image, followed by 256x256 and we even used a 512x512 image to test the output data. Regarding processing time, when using a 128x128 image there is no distinguishable delay introduced by the process. When using a bigger size as the input, there



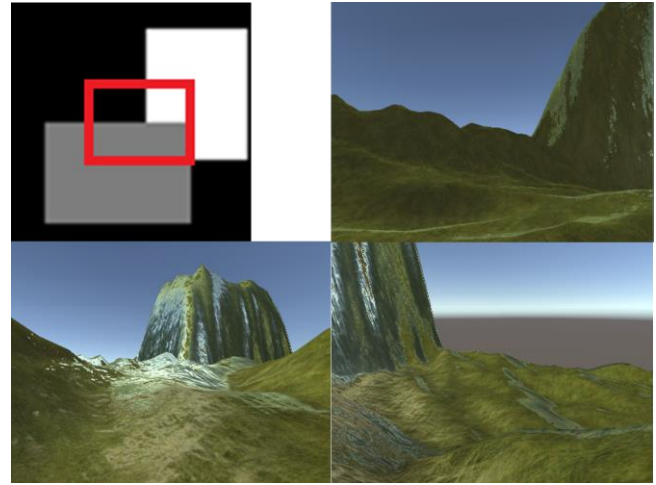
**Figure 1. Terrain generation using a pre-refined heightmap as a source**

is a slightly delay when employing a 256x256 image size, while using a 512x512 image size would bring a delay equal to 10-15 seconds at most. The results are acceptable as this part represents a pre-processing stage for the real-time generating method.

The second aspect when assessing the application's performances is to experiment various sizes of a chunk. In the current version of the application there is a 64x64 dimension size, so each time user reaches a certain distance, a new chunk of this size is generated in the direction of the user's movement. In this current implementation, it runs smoothly and without a visible delay when generating new chunks of terrain. Growing the dimension of the kernel used will definitely affect the run time of the algorithm. A size of 128x128 or 256x256 would not bring catastrophic results but thinking of a 1024x1024 or even bigger size would definitely introduce a delay that would affect the user's experience.

## CONCLUSION

As mentioned before, the purpose of this paper was to describe the functionality of an application that aims to create procedural terrain in real-time, being guided by the user's input. The user's input has no restriction in what concerns the level of detail, so a preprocessing algorithm had to be applied in order to obtain a viable height map used to create the mesh that would be rendered later. The real-time generation algorithm works by splitting the mesh into chunks and processing and displaying only those around the user's position. Future developments include porting the real-time computation to the GPU in order to improve performance.



**Figure 2. Terrain generation using a crude bitmap as a source**

## REFERENCES

1. Jacob Olsen. Realtime procedural terrain generation - realtime synthesis of eroded fractal terrain for use in computer games, (2004).
2. Jean-David G enevaux,  eric Galin, Eric Gu erin, Adrien Peytavie, and Bedrich Benes. 2013. Terrain generation using procedural models based on hydrology. *ACM Trans. Graph.* 32, 4, Article 143 (July 2013), 13 pages.
3. Alexandre Philippe Mangra, Adrian Sabou, Dorian Gorgan. Terrain Synthesis from Crude Heightmaps. *RoCHI 2016*, 113-118
4. H. Zhou, J. Sun, G. Turk and J. M. Rehg. Terrain Synthesis from Digital Elevation Models. *IEEE Transactions on Visualization and Computer Graphics* 13, 4 (2007), 834-848.
5. J. Schneider, T. Boldte, R. Westermann. Real-Time Editing, Synthesis, and Rendering of Infinite Landscapes on GPUs. *Conference on Vision, Modeling, and Visualization (VMV)*, (2006)
6. S. Worley. A Cellular Texture Basis Function. *SIGGRAPH '96 Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, (1996)
7. K. Perlin. Making Noise. (2002). [www.noisemachine.com/talk1/](http://www.noisemachine.com/talk1/)
8. R. Fisher, S. Perkins, A. Walker and E. Wolfart. Spatial filters - Median filter. (2003) <http://homepages.inf.ed.ac.uk/rbf/HIPR2/median.htm>
9. R. Fisher, S. Perkins, A. Walker and E. Wolfart. Spatial filters - Mean filter. (2003) <http://homepages.inf.ed.ac.uk/rbf/HIPR2/mean.htm>



# Improving an NP-Chunker by Exploring the Inner NP Dependency Structure

Mihaela Colhon

University of Craiova, Department of  
Computer Science, 13 Alexandru Ioan  
Cuza St.  
mcolhon@inf.ucv.ro

Dan Cristea

“Alexandru Ioan Cuza” University of Iași, Faculty  
of Computer Science, 16 Berthelot St. Iași  
Institute of Computer Science, the Iași branch of  
the Romanian Academy, 2 Codrescu St.  
dcristea@info.uaic.ro

## ABSTRACT

In this article we extend our previous work [3] dedicated to developing an automatic method for generating and applying syntactic patterns intended to recognize noun phrases. The intended scope of our studies is to automatically detect dependency relations within noun phrases exploiting some syntactic information of the phrases' words. The patterns are extracted from a corpus that has been automatically annotated for token and sentence limits, part of speech and noun phrases and manually annotated for dependency relations between words. The patterns have been generalized in order to cover more instances that were present in the corpus and to reduce the size of the grammar. In this paper we refine the syntactic patterns by considering not only the two terms of the dependency relation but also the terms surrounding these two that will be treated as contexts: left, middle, right depending on their position from the dependency terms.

## Author Keywords

Natural Language Processing; Dependency Parsing; HCI

## ACM Classification Keywords

H.5.2 User Interfaces (e.g., HCI): Natural language

## General Terms

Human Factors; Design; Measurement.

## INTRODUCTION

A way to incorporate into user interfaces knowledge expressed in natural language or modalities for communicating in a human to human style can be achieved by implementing Natural Language Processing tasks. In the context of Human Computer Interaction (shortly, HCI) the Natural Language Processing (shortly, NLP) issues range from various speech recognition systems, to natural language interfaces for various applications, or a multitude of machine translation systems.

Natural language processing applications, such as information retrieval, automatic translation, sentiment analysis or applications that automatically have to answer questions, require both semantic analysis and morpho-syntactic analysis of texts at different levels.

Particularly interesting from the HCI perspective are the research conducted in the domain of linguistics or the works put into constructing treebanks, both monolingual and parallel [2].

As an exemplification of the necessity of reliable natural language processing tools, let us consider the following question-answer example:

the information: "[...] Prices of sunflower oil rose by 27% due to low production [...]"

and the question: "For which food the price has increased by 27%?"

The answer "sunflower oil" is within the nominal group "Prices of sunflower oil", so the correct identification of the nominal group as well as the disambiguation of its structure is absolutely necessary in order to provide the correct answer. If the parser could not recover the structure of the nominal group then the correct answer candidate could not be found even if the phrase segment containing one or more nominal groups was been correctly identified. This problem also affects anaphora resolution systems, automatic syntactic analysis, but also automatic translation using syntactic trees [4, 16].

Automatic syntactic parsing of a text is used either singularly or in some linguistic processing chains in order to ease their purpose and to enhance precision, such as: coreference chains, entity recognition, nominal group processing (study in which the present work is included), or automated translation of a text. Syntactic parsing is the next level following morphological labelling, the latter being at the base of any processing chain.

Two of the most used formalisms to describe syntactic structures are in terms of hierarchies of constituents and dependency relations. Contiguous sequences of words are grouped under non-terminal symbols (as constituents), part of context free grammars, while dependency relations [11] are asymmetrical functional relations between pairs of words, considered head and modifier.

While these two traditions have sometimes been presented as competing with each other, a straightforward correspondence between a projective dependency analysis

(in which there are no crossing links) and a constituent structure analysis seems to exist [1].

Reliable dependency parsing is a notorious difficult problem in Natural Language Processing [5]. We describe in this paper a pattern-based approach in dependency parsing that addresses only Nominal Phrases (NPs).

To build a dependency treebank, the human experts must decide for each word which is the one it depends on. However, rather often there is no consensus among annotators on what the correct dependency structure for a particular sentence should be, because the decision regarding dependencies involve a deep interpretation process.

In contrast, in building phrase-structures, human annotators are confronted to much less ambiguity. This is because only a sequence of constituents should be indicated in the right hand side of a grammar rule, not also their relative roles/functions with respect to the parent constituent, as indicated by the left symbol. It is therefore normal that the effort of establishing correct dependency structures be paid back, and the difference stays in a much closer resemblance of a dependency structure to a semantic interpretation than in the case of a constituent structure. This aspect is even more important in the case of languages that have a free word order (as are for instance Czech, Romanian, etc.), where dependency treebanks are preferred to constituent structure representations (see, for instance the Prague Dependency Treebank [8]).

Using treebank data for training and evaluation of parsing systems is identified under the name of treebank parsing, a methodology that has been used to construct robust and efficient parsers for several languages over the last ten years [10]. For this kind of parsing, the treebank data is used to train the parser but also to evaluate the quality of the resulted parser with respect to accuracy as well as efficiency. Calacean and Nivre [3] report results on a MaltParser-based dependency [12] for Romanian, trained on a manually annotated Romanian Treebank<sup>1</sup> built in the RORIC-LING project [9]. Their precision for recognizing labelled relations are between 60.8% and 95.9%, depending on the length of the link, while the recall is in the range 71.3% - 96.3%. Their corpus includes only short sentences (with an average of 8.94 tokens per sentence) and a gold standard part-of-speech annotation.

In this paper we present an on-going research started in 2014 on a treebank dependency parsing mechanism that is restricted to NP chunks. The presented study refines the syntactic patterns defined in [5] by considering not only the two terms of the dependency relation but also the terms surrounding these two that will be treated as contexts: left, middle, right depending on their position from the

---

<sup>1</sup> <http://www.phobos.ro/roric/texts/xml/>

dependency terms. We work on Romanian, but the method is general enough to be applicable to other languages as well.

The paper is organized as follows: in Section 2 we present the Treebank corpus we use for our study. Section 3 describes the manner in which the collection of patterns extracted from an automatic annotation to NP chunks over the Treebank are organized. Section 4 describes the method and the results while Section 5 formulates a number of conclusions.

## THE TREEBANK

At <anonym-university>, a dependency Treebank for the Romanian language was built. We have used this resource in the training and evaluation stages of our proposed mechanism.

The corpus contains Romanian texts selected from a wide range of genres/registers of language<sup>2</sup>. Three levels of annotation have been added to the raw text and encoded in XML, by adopting a simplified form of the XCES standard [7]:

- Level-1: segmentation and lexical information. Sentences have their boundaries manually marked and each token has attached its part of speech, lemma, and morpho-syntactic information, by running an automatic processing chain that includes: tokenisation, POS-tagging and lemmatisation [14];
- Level-2: noun phrases. By exploiting Level-1 information, an NP-chunker [15] adds information regarding noun phrase boundaries and their head words;
- Level-3: syntactic dependency data. During a manual annotation phase [13] each token of all sentences has been complemented with its head-word and the dependency relation towards the head.

The Treebank thus acquired contains 2,630 sentences, in which 7674 NP structures were identified by the NP chunker.

## THE DATABASE OF NOUN PHRASES PATTERNS

The primary data extracted from the corpus is used to associate dependency structures with each sequence of MSD<sup>3</sup> tags corresponding to NPs extracted from the corpus,

---

<sup>2</sup> The corpus mainly includes Romanian translations of the first chapter of George Orwell's novel "1984", Romanian parts from the JRC-Acquis corpus, Romanian Wikipedia texts, grammar texts used in high schools, etc.

<sup>3</sup> Morpho Syntactic Description – notation used in the Multext projects [6].

which we will call in the following morphological structures.

The corpus includes only contiguous NPs. As we have already said, the corpus used in this study puts in evidence three levels of annotations: POS tags, NP chunks and dependency relations. For each NP chunk, we extracted the morphological structure and the configuration of dependency relations manually marked among the words of the NP chunk.

To syntactic constituents of the sentence correspond dependency structures organized in (sub)trees. Each node of such a (sub)tree is a word, connected with a dependency relation to its head word. The roots of these (sub)trees are the only elements related to words outside the constituents themselves. The only exception is the very root of the sentence which is not related to another word. In the same way, the head word of an NP is the only word belonging to the NP related outside the NP.

For example, if we take the following bracket representation for a Romanian noun phrase „*lamele de ras tocite*” (En. “*the blunt razor blades*”):

```
[NP [Ncfpry lamele] [Spsa de] [Ncfsry ras] [Afpfp-n tocite]]
```

its morphological structure is:

```
Ncfpry Spsa Ncfsry Afpfp-n4
```

and its internal dependencies are:

```
a.subst.(lamele-1,de-2)
prep.(de-2,ras-3)
a.adj.(lamele-1,tocite-4)
```

In the notations above the name of relation is placed in front of a pair of words, the first one being the head and the second - the modifier word. The number attached to a word denotes its position inside the chunk.

A database table is built out of the 3 layers of notations in the corpus: each record in this table includes a triplet

```
<seq-msd> <seq-rels> <seq-heads>
```

in which <seq-msd> represents a morphological structure (a pattern of MSD tags), <seq-rels> represents the sequence of dependency relations (transferred from the words to the MSD representations on the respective positions - the first position is counted as 1) and <seq-heads> is the sequence of head positions in the pattern<sup>5</sup>. The 0-rel always marks in <seq-rels> the relation of the head word of the NP going out of the NP, and with 0-

<sup>4</sup> See the Appendix for the meaning of the MSD tags in this paper.

<sup>5</sup> Of course, the property of unique occurrence of a record in the database is observed here as well.

head is noted in <seq-heads> the exterior head of the head word of the NP. The head of an NP chunk is usually attached, in the syntactic tree of the sentence, to another word which, of course, is not part of that NP sequence.

For example, to the NP “*lamele de ras tocite*”, the following entry corresponds in the database:

```
<Ncfpry Spsa Ncfsry Afpfp-n>
<0-rel a.subst. prep. a.adj.>
<0-head 1 2 1>
```

Based on the dependency data coded in the corpus the entries of the database are grouped in three categories:

- with no external heads: no 0-rel in the <seq-rels> field;
- with one external head: exactly one 0-rel in the <seq-rels> field;
- with more than one external head: minim two 0-rel in the <seq-rels> fields.

Using these representations one can easily detect the incorrect NP sequences marked in the corpus but also the ambiguous <seq-msd> sequences with respect to their dependency structures.

Indeed, exploiting this grouping, the NP chunks incorrectly marked in the corpus during the automatic NP-chunking results immediately as they correspond to <seq-msd> entries which include more than one 0-rel in the <seq-rels> sequence.

Opposite to this case, <seq-msd> sequences with one single external head represent correct NP chunks that are connected with the other words of the sentences they belong to by means of their heads.

Correct MSD sequences of NP chunks are also the <seq-msd> entries in the database with no 0-rel. These cases correspond to verb elliptical sentences where the NP is the very root of the sentence - the main verb is missing.

The ambiguous dependency constructions popup immediately as being the sets of entries in the database that have the same <seq-msd> sequence.

## APPLYING PATTERNS

Our method for generating and applying corpus based patterns in dependency parsing works as follows (see Figure 1): starting from flat NP sequences and using a set of syntactic patterns extracted from the training corpus, we identify dependency links between the words of the NPs, based on their MSD finger-print. By this, the flat NP sequences become dependency sub-trees. In order to do that, the dependency relations, considered independent one of the others, are decoupled from the particular morphological structures they occur in. The set of contexts of each relation is then tried to be generalized. The aim of

the generalization is to reduce the number of dependency relations patterns extracted from the corpus, but also to infer deducible sequences not instantiated in the corpus. During the pattern-generalization process, the following steps are repeated for all records of the database in which a relation  $R$  occurs between identical tags:

for each `<seqrels>` sequences in the database, three types of contexts are marked:

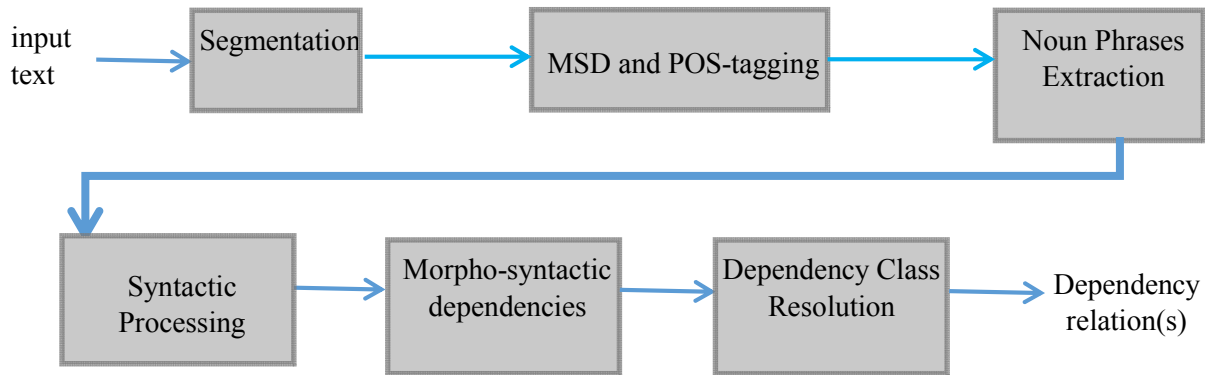
- **left context:** is represented by the sequence of MSD tags in `<seq-msd>` appearing to the left of the position of the first element involved in the dependency  $R$ ;
- **middle context:** is represented by the sequence of MSD tags appearing in between the two tags involved in the targeted dependency relation

- **right context:** is represented by the sequence of MSD tags appearing to the right of the second element involved in the targeted dependency relation  $R$ .

For each dependency relation, the dependents are represented by two MSD tags:

- 1: the MSD tag for the head
- 2: the modifier MSD tag

The contexts (if they are present in the pattern structure) may be optional or mandatory: the optional ones are marked with `?` and the mandatory ones with `{1}`.



**Figure 1. Dependency Parsing Mechanism within Noun Phrases**

Let us now take three examples of NPs:

NP#1: “*Fetele acestea frumoase*” (En: “*These beautiful girls*”)

NP#2: “*Mâinile lui curate*” (En: “*His clean hands*”)

NP#3: “*Lumânările aprinse*” (En: “*candles lit*”)

In all these NPs the same relation (adjectival attribute, noted with “*a.adj.*”) is found between words displaying identical MSD tags (that is `Ncfpry` as head and `Afpfp-n` as modifier). The following patterns for this dependency relation are found:

From NP#1:

`<1:Ncfpry Dd3fpr- 2:Afpfp-n>`

From NP#2:

`<1:Ncfpry Pp3mso- 2:Afpfp-n>`

From NP#3:

`<1:Ncfpry 2:Afpfp-n>`

After the generalization process, all these representations will be merged into a single one with an optional middle context:

`<1:Ncfpry (Dd3fpr-|Pp3mso)? 2:Afpfpn>`

The middle context in the generalized pattern is optional because it can either contain one of the tags separated by the “or” operator (“|”) or can be an empty one like in the last example.

After generalization, the patterns extracted from the corpus were grouped into 24 sets, each corresponding to one relation and covering the span of an entire NP.

The evaluation scores obtained are given in the next section.

## EVALUATION

The total Treebank sentences were split into a training set and a testing set. The parser was trained on approximately 90% of the Treebank and evaluated on the remaining 10% using a 10-fold cross-validations policy, which guaranteed no intersection between training and evaluation sentences.

For the testing set, the most frequent relations identified by the parser were:

- the adjectival attribute (*a.adj.*),
- nominal attribute (*a.subst.*),
- determiner (*det.*) and
- preposition (*prep.*).

No restrictions of projectivity of the generated dependency structures have been included at this moment. The obtained scores are given in Table 1. As one can observe, the best accuracy was obtained for determiner relation while the lower was for the adjectival attribute relation. Nevertheless, the evaluation scores give hope for continuing our pattern-based approach in Romanian dependency parsing.

Accuracy Dependency relations	Precision	Recall	F-measure
<b>ALL</b>	0.63	0.74	0.68
<b>a.adj.</b>	0.54	0.86	0.66
<b>a.subst.</b>	0.67	0.72	0.69
<b>det.</b>	0.90	0.82	0.86
<b>prep.</b>	0.74	0.66	0.70
<b>Unlabelled relations</b>	0.72	0.78	0.75

**Table 1: Evaluation scores**

## CONCLUSION

It is well-known that there is a difficult trade in designing proper generalization patterns, because making them too lax could trigger false instances. On the other hand, making them too straight will imply low recall scores. There is perhaps more to be done in this direction.

We are aware that the model would gain in precision if lexical information would be included, by enriching the MSD tags of the generated patterns with lemmas. In our experiments till now we excluded lexical information, because lexical information presupposes a much larger training corpus, which was not at our fingers during this phase of research.

## REFERENCES

1. Michael Brody. 1994. Phrase structure and dependence. Working papers in the theory of grammar. 1(1), Theoretical Linguistics Programme, Budapest Univ.
2. Marian Cristian Mihăescu, Mihaela Gabriela Țacu, Dumitru Dan Burdescu. 2014. Use Case of Cognitive and HCI Analysis for an E-Learning Tool, *Informatica* 38: 273-279
3. Mihaela Călăcean, Joakim Nivre. 2009. A Data-Driven Dependency Parser for Romanian, *Proceedings of TLT-7*.
4. Alexandra Cristina Cristea. 2014. Syntactic Study on Nominal Groups in Romanian (in Romanian), Dissertation Thesis, “Alexandru Ioan Cuza” University of Iași
5. Mihaela Colhon, Dan Cristea. 2014. Automatic Extraction of Syntactic Patterns for Dependency Parsing in Noun Phrase Chunks, *Bucharest Working Papers in Linguistics*, vol. XVI, nr. 1, ISSN 1454-9328
6. Tomaž Erjavec. 2010. Multext-east version 4: Multilingual morphosyntactic specifications, lexicons and corpora. In *LREC*.
7. Nancy Ide, P. Bonhomme, L. Romary. 2000. Xces: An xml-based encoding standard for linguistic corpora. In *Proceedings of the Second International Language Resources and Evaluation Conference*. Paris: European Language Resources Association.
8. Jan Hajič, Alena Böhmová, Eva Hajičová, Barbora Vidová-Hladká. 2000. The prague dependency treebank: A three-level annotation scenario, A. Abeillé, editor, *Trebanks: Building and Using Parsed Corpora*, Amsterdam:Kluwer: 103-127.
9. Florentina Hristea, Marius Popescu (eds.). 2003. *Building Awareness in Language Technology*, University of Bucharest Publishing House.
10. Svetoslav Marinov, Joakim Nivre. 2005. A Data-Driven Dependency Parser for Bulgarian. In *Proceedings of TLT 2005*: 89-100.
11. Igor Mel'čuk. 1987. *Dependency Theory: Syntax and Practice*, Albany, NY: SUNY Press
12. Joakim Nivre, Johan Hall, Jens Nilsson. 2006. MaltParser: A data-driven parser-generator for dependency parsing. *Proceedings of LREC-2006*: 2216-2219
13. Ceneș A. Perez. 2012. Casuistry of Romanian functional dependency grammar, M. A. Moruz, D. Cristea, D. Tufis, A. Iftene, H. N. Teodorescu (eds.) *Proceedings of the 8th International Conference “Linguistic Resources And Tools For Processing Of The Romanian Language”*: 19-28
14. Radu Simionescu. 2012. Graphical grammar studio as a constraint grammar solution for part of speech tagging, M. A. Moruz, D. Cristea, D. Tufis, A. Iftene, H. N. Teodorescu (eds.) *Proceedings of the 8th International Conference “Linguistic Resources And Tools For Processing Of The Romanian Language”*: 109-118.



15. Radu Simionescu. 2012. Romanian deep noun phrase chunking using graphical grammar studio, M. A. Moruz, D. Cristea, D. Tufis, A. Iftene, H. N. Teodorescu (eds.) Proceedings of the 8th International Conference ‘Linguistic Resources And Tools For Processing Of The Romanian Language’’: 135-143.
16. David Vadas, James R. Curran. 2011. Parsing Noun Phrases in the Penn Treebank, School of Information Technologies, University of Sydney, Australia.

**APPENDIX**

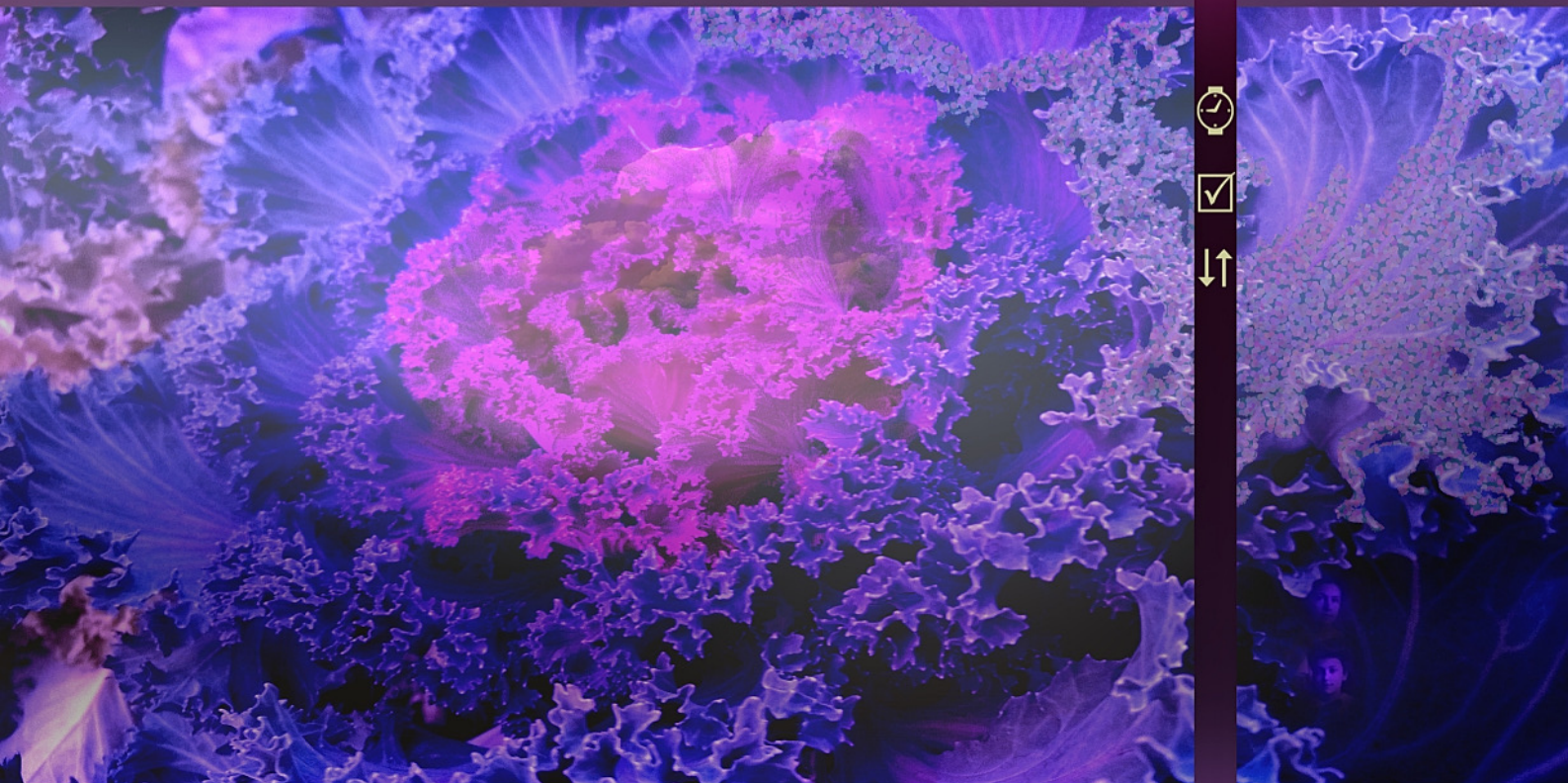
The following table gives the Morpho-syntactic descriptions (shortly, MSD) used in this paper.

Afpfp-n	Adjective qualifier positive feminine plural -definiteness
Dd3fpr-	Determiner demonstrative third feminine plural direct
Ncfpry	Noun common feminine plural direct +definiteness
Ncfsry	Noun common feminine singular direct +definiteness
Pp3mso-	Pronoun personal third masculine singular oblique
Spsa	Adposition preposition simple accusative

MSD tag	The meaning of the notation (according to MULTEXT-East lexical specifications)
---------	--



# RoCHI – International Conference on Human-Computer Interaction



**Proceedings of the 14<sup>th</sup> Conference  
on Human-Computer Interaction – RoCHI 2017**  
11—12 September 2017 • Craiova, Romania

**RoCHI** is the local group of ACM SIGCHI – Special Interest Group in Computer-Human Interaction in Romania, and is intended to serve as an inter-disciplinary forum for the exchange of ideas and experience in the HCI field.

[rochi.ici.ro](http://rochi.ici.ro)  
[rochi.utcluj.ro](http://rochi.utcluj.ro)

ISSN 2501-9422  
ISSN-L 2501-9422

