VOLUME 7 NUMBER 1 BANJA LUKA, JUNE 2017 (1-56) ISSN 2232-9625 (Print) ISSN 2233-0194 (Online) UDC 004

# Journal of Information Technology and Applications

(BANJA LUKA)

JITA

Exchange of Information and Knowledge in Research

APEIRON

THE AIM AND SCOPE

The aim and scope of the Journal of Information Technology and Applications (JITA) is:

- to provide international dissemination of contributions in field of Information Technology,
- to promote exchange of information and knowledge in research work and
- to explore the new developments and inventions related to the use of Information Technology towards the structuring of an Information Society.

JITA provides a medium for exchanging research results and achievements accomplished by the scientific community from academia and industry.

By the decision of the Ministry of Education and Culture of the Republic of Srpska, no.: 07.030-053-160-4/10 from 3/3/2010, the journal „Journal of Information Technology and Applications“ Banja Luka is registered in the Registry of public organs under the number 591.
Printed by Markos, Banja Luka in 300 copies two times a year.

**Indexed in:** LICENSE AGREEMENT, 3.22.12. **EBSCO** Publishing Inc., Current Abstracts

ebscohost.com
indexcopernicus.com
citefactor.org/contact
scholar.google.com
doisrpska.nub.rs
crossref.org
road.issn.org
citefactor.org
cosmosimpactfactor.com

Printed on acid-free paper

Full-text available free of charge at http://www.jita-au.com

UDC 004

ISSN 2232-9625 (Print)

ISSN 2233-0194 (Online)

JITA

VOLUME 7 NUMBER 1 BANJA LUKA, JUN 2017 (1-56)

## Contents

The content of this issue of JITA consists of six papers. The first paper, entitled "Use of computer search algorithms in the research of statistical, semantic and contextual rules of language in digital information space" by Z. Avramovic, D. Marinkovic, I. Lastric Igor, and G. Radic, presents the results of researching the interdependence between information technology and linguistics in the modern information society. The aim of this paper is to extend knowledge about the possibilities of application of information technologies in researching rules of language, as well as emphasizing the importance that language technologies have in the field of linguistic research, preservation of language and culture and national identity.

The next paper is "Space Complexity Analysis of the Binary Tree Roll Algorithm" by A. Božinovski et al. This paper presents the space complexity analysis of the binary tree roll algorithm. The space complexity is analyzed theoretically and the results are then confirmed empirically. The space complexity is shown, both theoretically and empirically, to be logarithmic in the best case and linear in the worst case, whereas its average case is shown to be dominantly logarithmic.

The third article "The use of digital signature in electronic communication in BiH - Research" by T. Talić describes the results of usage of digital signature in Internet communication, having in mind that the increasing use of electronic mail for identity theft and unsolicited marketing reduced the credibility of email as a communication tool. Authentication of the sender is well known defense against such attacks and one of the methods to ensure secure communication via e-mail is the use of digital signature.

In the paper "Expert systems in a cloud computing environment model for fast-paced decision making", M. Bakator and D. Radosav outline the use of cloud computing technologies in expert systems implementation. The proposed cloud-based expert system offers effective, fast and reliable support for individuals or organizations when it comes to fast-paced decision making.

"Modeling the process of IS auditing in the public administration using UML diagrams" by D. Drljača, B. Latinović, and D. Starčević proposes a model for auditing of information systems in the public administration institutions. The model intends to explain the audit process using a visual representation of the process with UML diagrams.

The last article in this issue "Comparison of perceived interactivity measures of actual websites interactivity" by V. Štavljanin and M. Jevremović, investigates the importance of the interactivity for digital marketing. Paper reports a study in which two websites of low and high interactivity were developed and in an experimental setting as stimuli used to test three perceived interactivity measures. Results show that all measures estimated perceived interactivity of a high interactivity website better than of a low interactivity website.

On behalf of the Editorial Board we would like to thank the authors for their high-quality contributions, and also the reviewers for the effort and time invested into the preparation of this issue of the Journal of Information Technology and Applications.

# Use of Computer Search Algorithms in the Research of Statistical, Semantic and Contextual Rules of Language in Digital Information Space

**Zoran Ž. Avramović, Dražen Marinković, Igor Lastrić**

*Pan-European University Apeiron, Banja Luka, the Republic of Srpska, Bosnia and Herzegovina*

*zoran.z.avramovic@apeiron-edu.eu*

*drazen.m.marinkovic@apeiron-edu.eu*

*igor.t.lastric@apeiron-edu.eu*

**A general survey**

DOI: **10.7251/JIT1701005A**

UDC: **004.7:81'33**

**Abstract:** This paper will discuss and practically explore the interdependence between information technology and linguistics in the modern information society.

The relationship between information technology and linguistics, which has opened new opportunities in linguistic research, will be practically seen in the application of linguistic engineering in researching rules of language.

The aim of this paper is to extend knowledge about the possibilities of application of information technologies in researching rules of language, as well as emphasizing the importance that language technologies have in the field of linguistic research, preservation of language and culture and national identity.

**Keywords:** information technology, search algorithm, rules of language, linguistic engineering, digital information space.

## Introduction

This paper explores and shows rules of language through finding linguistic data in the digital information space, using different methods based on the originally developed, programmed search algorithms in procedural and object-oriented languages that use the system to manage databases and relational model for final realization.

The subject of this paper is to explore the methods of finding, processing and presentation of data in digital libraries and the possibilities of improving the quality of that finding, by finding a faster search query response.

Comparing different methods of information retrieval in digital libraries, the paper starts from the basic, traditional methods of finding information and determining their effectiveness.

The objective of this paper is to determine whether and how advanced software methods help in finding quality linguistic data and their analysis.

The novel "The Bridge on the Drina" by the Nobel Prize winner Ivo Andric (Serbian language, Cyrillic, Ekavian dialect) is taken as the basis for the digital object record on which the research and the analysis of search results and extraction of data, information and knowledge will be carried out. [3]

## Algorithms, Languages and Search Programs

An algorithm is a list of steps to follow in order to solve a problem, without ambiguity and vague-

ness. It relates to the principles of determinacy and finality. In computing environment, algorithms are implemented in a programming language.

In computer science, a search algorithm is an algorithm that retrieves information stored within some data structure.

The basic search algorithms are simple to implement and are suitable for search of static data sets (regulated and unregulated table). This group includes: sequential search and binary search.

Sequential search algorithm is the simplest to implement, but its performance is not so good. It starts to check each item in order to determine whether that item is the one we are looking for. If so, the search is successfully completed, otherwise it goes on, until it finds the desired item or until it examines all items.

In preparing this paper, we have primarily used Visual Basic, a programming language intuitively designed, based on events as an integral part of the programming system from Microsoft, designed to enable rapid application development - RAD with a graphical user interface - GUI and to communicate with databases (such as the DAO, RDO or ADO) and creating ActiveX controls and objects.

In our advanced analysis, processing and research, general purpose programming language C++ has found its application. It is an object-oriented language which provides facilities for low-level memory manipulation. It is designed to support system programming, especially in the case of limited resources, but it is widely used with large systems designed for efficient and flexible operation.

C++ programming language was initially standardized in 1998 by the International Organization for Standardization (ISO). Before the initial standardization, C++ was developed by Bjarne Stroustrup, as an extension of the C language as he wanted an efficient and flexible language similar to C. In this paper, the latest standard version, ratified and published by ISO in December 2014 as ISO/IEC 14882:2014 (informally known as C++14), has found its application.

## Ivo Andric - "The Bridge on the Drina"

"The Bridge on the Drina" is a historical novel by Ivo Andric, which, among other literary works, won the Nobel Prize for Literature in 1961.

The novel tells the story of the bridge, Visegrad and human destinies associated with it and spans about four centuries. The bridge stands as a silent witness to history from its construction by the Ottomans in the mid-16th century until its partial destruction during World War I. The bridge is an object around which the destinies of Visegrad inhabitants intertwine, of Muslims, Orthodox Christians, Jews and immigrant Catholics, who play dramatic roles in a large theater of history.

Ivo Andric wrote this novel during World War II when he moved to Belgrade in 1944, after he had worked as an ambassador in Berlin.

Inspiration for the novel Andric found in his own life - he spent his childhood in Visegrad with his aunt who had raised him after his mother lost revenue when his father died. He finished primary school in Visegrad where he was looking at the impressive bridge on the Drina.

The Serbian language in Andric's novel is on one hand seemingly simple and easy to understand, but on the other hand we can see that every word in it is meticulously measured and harmoniously blended. Many think the language is so special, calling it Andric's – Ekavian dialect with syntax which is sometimes characteristic of English language as well as of Bosnian language. In fact, this is the language that Andric learned in his childhood, the true language of that region.

In this literary work we find many examples of folk wisdom in stories, poems and legends as well as in customs and beliefs. Significant attention is given to descriptions, both external and psychological.

Apparently, the novel does not have a theme and strict storyline, something that would link the story with other stories, following one after another, but there is the bridge on the Drina and it is the binding element. It symbolizes the strength and permanence, continuity and consistency despite all the disasters that threaten it. All inhabitants of Visegrad and the surrounding area are connected to this bridge. Compared to it, life expectancy is short and insignificant, so the writer points it out only in carefully chosen moments of human misery.

## Search Results Display

The analysis of the text in the novel "The Bridge on the Drina" by Ivo Andric (without title and text

on the cover and inner cover page) showed that the total number of characters in it is 532,109 and the following is determined (Table 1):

***Table 1.*** *Evidence-based analysis of the novel*

| No. | What | How much |
|---|---|---|
| 1 | Total number of characters in the novel | 514,558 |
| 2 | Total number of punctuation | 17,448 |
| 3 | Total number of numeric digits | 103 |
| 4 | Total number of words in the novel | 115,395 |
| 5 | Total number of sentences in the novel | 6,011 |

Study of vowel representation is shown in Table 2 and Figure 1

***Table 2.*** *Number and frequency of occurrence of vowels in the novel*

| Vowel (uppercase and lowercase letter) | Number of occurrences | Percentage of the vowels | Percentage of the most present | Relative frequency of all letters |
|---|---|---|---|---|
| **A** | 62,837 | 26.668 | 100.000 | 3.66 |
| **I** | 51,404 | 21.816 | 81.805 | 3.00 |
| **O** | 50,985 | 21.638 | 81.139 | 2.97 |
| **E** | 47,547 | 20.179 | 75.667 | 2.77 |
| **U** | 22,855 | 9.699 | 36.372 | 1.33 |
| **Total** | **235,628** | **100.000** | | |

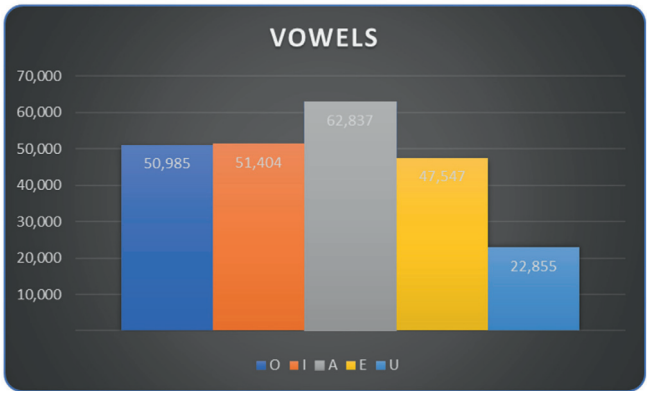



***Figure 1.*** *Histogram of distribution of the total number of vowels in the novel*

The number and relative abundance of uppercase and lowercase letters among vowels are given in Table 3.

***Table 3.*** *Number and relative abundance of uppercase and lowercase letters among vowels*

| No. | Vowel | As uppercase letter | As lowercase letter | Letter case percentage ratio |
|---|---|---|---|---|
| 1 | **A** | 827 | 62,010 | 1.333 |
| 2 | **I** | 554 | 50,850 | 1.089 |
| 3 | **O** | 525 | 50,460 | 1.040 |
| 4 | **E** | 51 | 47,496 | 0.107 |
| 5 | **U** | 340 | 22,515 | 1.510 |

It is noted that the vowel "e" rarely appears at the beginning of a sentence and/or at the beginning of proper nouns.

The research of the word length in the novel according to number of letters has given interesting results (Table 4 and Figure 2).

***Table 4.*** *Number of words in the novel according to number of letters*

| Word length | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| Number of occurrences | 12,473 | 22,298 | 10,804 | 16,905 | 14,729 | 13,719 | 9,506 |

| Word length | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|
| Number of occurrences | 6,664 | 4,162 | 2,201 | 1,212 | 475 | 201 | 68 | 41 |

| Word length | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
|---|---|---|---|---|---|---|---|---|
| Number of occurrences | 16 | 6 | 4 | 3 | 1 | 0 | 0 | 1 |

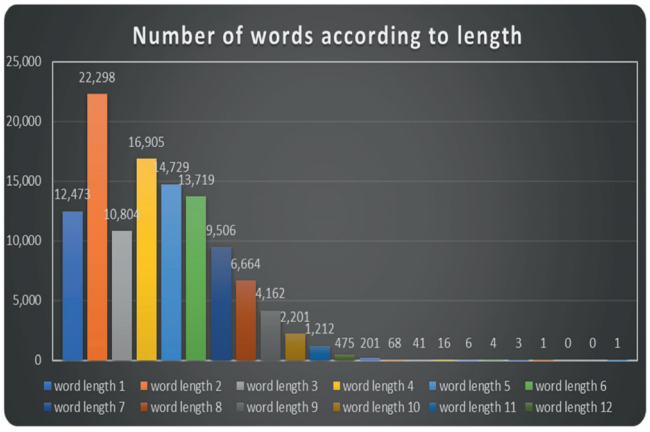



***Figure 2.*** *Histogram of number of words in the novel according to number of letters*

Computer analysis of the text in the novel "The Bridge on the Drina" by Ivo Andric has found that the average length of sentences is 19.20 words.

In this literary work, the average number of letters in the sentence is 85.61.

The longest word in observed and analyzed text has 23 letters.

In this paper we have also researched the presence of the consonants in the novel. For example, in alphabetical order, the first two of them appears: B - 7,881 and V - 19,380 times.

## CONCLUSION

Based on the conducted research and analysis, we have concluded that the advanced search methods help in finding better data quality and thus the information. This was evident in comparing the efficiency and accuracy of finding the data with the use of advanced methods compared to conventional methods.

We can say that finding information using computational methods gives immeasurably better results compared to data retrieval that relies only on the classical methods. We want to emphasize how important is the presence and use of digital resources of analyzed works, especially morphological dictionaries (which are of great importance for morphologically rich languages like Serbian language).

The main focus of our paper was on computer search and finding data because digital libraries offer better and broader search options – through the full text of documents contained in them and also through metadata that describes documents. We believe that a digital library should be much more than a collection of documents available in digital form. We partly focused the subject of our research towards desire to determine whether the present digital libraries can respond to such requests.

We emphasize that the field of separation of concerns, processing and presentation of data, as a subfield of field of natural languages, largely depends on description of observed natural language. Though this field is quite advanced in some languages, such as English, we note that in Slavic languages, especially Serbian language, it is still relatively new.

## REFERENCES:

[1] Aleksandra S. Trtovac, Metadata descriptors and content descriptors for finding information in digital libraries, doctoral dissertation, University of Belgrade, Faculty of Philology, Belgrade, 2016.

[2] Stasa I. Vujicic Stankovic, Ontology-guided Information Extraction (model for Serbian language), doctoral dissertation, University of Belgrade, Faculty of Mathematics, Belgrade, 2016.

[3] Ivo Andric, The Bridge on the Drina, published by the Institute for Textbooks and Teaching Aids, Belgrade, 2012. (leter: Cyrillic, dialect: ekavian, number of pages: 376)

[4] Mladenovic, M. i Mitrovic, J. Semantic Networks for Serbian: New Functionalities of Developing and Maintaining a WordNet Tool. Natural Language Processing for Serbian: Resources and Applications (pages 1-11), Belgrade, University of Belgrade, 2014.

[5] Mitar Pesikan, Jovan Jerkovic, Mato Pizurica, Pravopis srpskoga jezika, Third Edition, Matica Srpska, Novi Sad, 2016.

[6] Donald Erwin Knuth, *The Art of Computer Programming Volume 3: Sorting and Searching third edition*, 1998, Addison-Wesley Professional. ISBN978-0-201-48541-7.

[7] Branimir Covic, Milos Crnjanski's Migrations in the context of modern Russian historical novel, Belgrade, 2001 (S-20287)

Submitted: May 25, 2017.
Accepted: June 7, 2017.

# Space Complexity Analysis of the Binary Tree Roll Algorithm

**Adrijan Božinovski[1], George Tanev[1], Biljana Stojčevska[1], Veno Pačovski[1], Nevena Ackovska[2]**

*[1]School of Computer Science and Information Technology, University American College Skopje, Macedonia*

*[2]Faculty of Computer Science and Engineering, University "Sv. Kiril i Metodij", Skopje, Macedonia*

*[1]bozinovski@uacs.edu.mk, [1]george.tanev@uacs.edu.mk, [1]stojcevska@uacs.edu.mk, [1]pachovski@uacs.edu.mk, [2]nevena.ackovska@finki.ukim.mk*

**Contribution to the state of the art**

DOI: **10.7251/JIT1701009B** UDC: **004.42:510.5**

**Abstract:** This paper presents the space complexity analysis of the Binary Tree Roll algorithm. The space complexity is analyzed theoretically and the results are then confirmed empirically. The theoretical analysis consists of determining the amount of memory occupied during the execution of the algorithm and deriving functions of it, in terms of the number of nodes of the tree $n$, for the worst - and best-case scenarios. The empirical analysis of the space complexity consists of measuring the maximum and minimum amounts of memory occupied during the execution of the algorithm, for all binary tree topologies with the given number of nodes. The space complexity is shown, both theoretically and empirically, to be logarithmic in the best case and linear in the worst case, whereas its average case is shown to be dominantly logarithmic.

**Keywords:** Binary Tree Roll Algorithm, space complexity, theoretical analysis, empirical analysis.

## Introduction

Trees are fundamental concepts in computer science, and are frequently used to keep track of ancestors or descendants, sports tournaments, organizational charts of large corporations and so on [19]. Trees are one of the basic data structures used in combinatorial algorithms [13], search techniques (e.g., [8, 4]), and game playing [17]. This paper also points out the use of binary trees for generating integer sequences, which are important in information forensics [20], cryptography [9], and security [12]. Binary trees have been shown to be very useful in mathematics and computer science and as such have been extensively studied. Several variations of the binary tree structure have been conceived, such as binary search trees, red-black trees [10], AVL trees [1], B-trees [3], and so on. Binary trees are often used as auxiliary data structures in other research endeavors, both practical (e.g., [18, 15]) and theoretical (e.g., [16, 14]), but occasionally are the subject of the research itself (e.g., [2]).

Binary Tree Roll is an operation by which all of the nodes of a binary tree are rearranged in such a way, so that two of the depth-first traversals of the newly obtained binary tree yield the same results as other two traversals of the original binary tree. The graphical representation of the newly obtained binary tree is that it appears to be rolled at a 90 degree angle (either counterclockwise or clockwise, depending on the direction of the applied roll operation) relative to the original binary tree; hence the name "Binary Tree Roll".

This operation was introduced and defined in [5]. There are two variants of the Binary Tree Roll Operation: a counterclockwise (CCW) and a clockwise (CW) roll. The counterclockwise roll of a binary tree, abbreviated as CCW(), is defined as follows. Given two binary trees $T_1$ and $T_2$, as well as their respective preorder(), inorder() and postorder() traversal functions, operation CCW() is defined as in Definition 1:

$$CCW(T_1) = T_2 \Leftrightarrow (preorder(T_1) = inorder(T_2) \wedge inorder(T_1) = postorder(T_2)) \quad (1)$$

In other words, upon CCW(), the preorder traversal of the original tree is identical to the inorder traversal of the tree obtained by the counterclockwise roll, and the inorder traversal of the original tree is identical to the postorder traversal of the tree obtained by the counterclockwise roll.

Likewise, the clockwise roll of a binary tree, abbreviated as CW(), is defined as in Definition 2:

$$CW(T_1) = T_2 \Leftrightarrow (inorder(T_1) = preorder(T_2) \wedge postorder(T_1) = inorder(T_2)) \quad (2)$$

Similarly, upon CW(), the inorder traversal of the original tree is identical to the preorder traversal of the tree obtained by the clockwise roll, and the postorder traversal of the original tree is identical to the inorder traversal of the tree obtained by the clockwise roll.

A graphical explanation was given in [5], showing how the resulting binary tree is obtained visually, so as to comply with definition (1) or (2), depending on the direction of the roll. The *downshift* visual operation, illustrated in Figure 1, was also presented. It was shown that CCW() and CW() are inverses of each other, and algorithms for CCW() and CW() were given, which did not require obtaining the traversals of the input tree in order to generate the rolled tree.

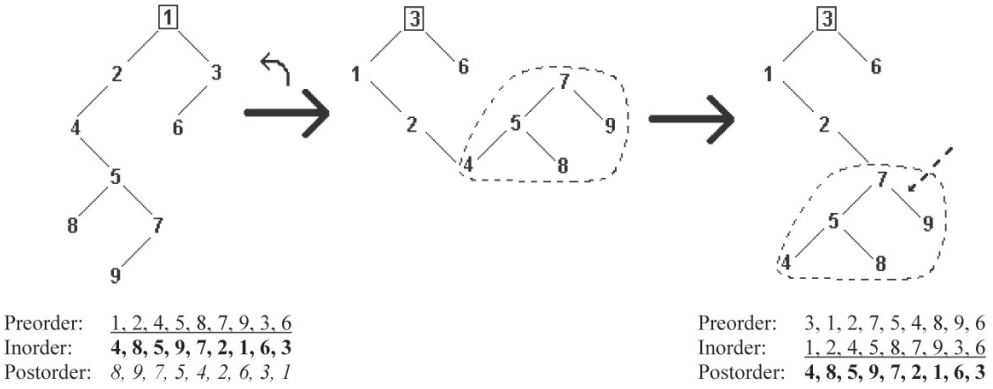



***Figure 1.*** *Graphical explanation of the CCW() algorithm, and an example of a downshift [5]*

Structurally, the algorithm presented in [5] contains a trivial case, two basic cases, and a third, more complex one. The pseudocode for both the CCW() and CW() variations of the algorithm are shown in Figure 2.

The algorithm takes two input parameters, which represent two binary tree nodes: the root of the tree to be processed, and its predecessor. The predecessor's initial value is always NULL, since the root of the input tree never has a predecessor node. However, the value of the predecessor parameter changes as further recursive calls to the algorithm are being invoked from inside the function itself. Moreover, the values of both the root and the predecessor nodes are guaranteed to change within subsequent recursive function calls, since the entire structure of the binary tree is rearranged after the roll operation executes fully.

The motivation for this paper was the fact that the binary tree roll algorithm, in either its CCW() or CW() variant, has so far been analyzed for time complexity [7] but not for space complexity. The latter is the goal of this paper and it will be done as follows, focusing on the CCW() variant. First, a theoretical analysis of the space complexity will be given, treating all cases of the algorithm execution. The principle of space com-

```
1.  CCW(&root, &predecessor)
2.  {
3.   if(root != NULL)
4.   {
5.    if(root.rSn == NULL)
6.    {
7.     root.rSn = root.lSn;
8.     root.lSn = NULL;
9.     CCW(root.rSn, root);
10.   }
11.   else
12.   {
13.   if(root.rSn.rSn == NULL)
14.   {
15.    root.rSn.rSn = root.rSn.lSn;
16.    root.rSn.lSn = root;
17.    root = root.rSn;
18.    root.lSn.rSn = root.lSn.lSn;
19.    root.lSn.lSn = NULL;
20.    if(predecessor != NULL)
21.          predecessor.rSn = root;
22.    CCW(root.lSn.rSn, root.lSn);
23.    CCW(root.rSn, root);
24.   }
25.   else
26.   {
27.    CCW(root.rSn, root);
28.    define leftmost = root.rSn;
29.    while(leftmost.lSn != NULL)
30.          leftmost = leftmost.lSn;
31.    leftmost.lSn = root;
32.    define newroot = root.rSn;
33.    root.rSn = NULL;
34.    root = newroot;
35.    if(predecessor != NULL)
36.         predecessor.rSn = root;
37.    CCW(leftmost.lSn, leftmost);
38.   }
39.  }
40.  }
41. }
```

a)

```
1.  CW(&root, &predecessor)
2.  {
3.   if(root != NULL)
4.   {
5.    if(root.lSn == NULL)
6.    {
7.     root.lSn = root.rSn;
8.     root.rSn = NULL;
9.     CW(root.lSn, root);
10.   }
11.   else
12.   {
13.   if(root.lSn.lSn == NULL)
14.   {
15.    root.lSn.lSn = root.lSn.rSn;
16.    root.lSn.rSn = root;
17.    root = root.lSn;
18.    root.rSn.lSn = root.rSn.rSn;
19.    root.rSn.rSn = NULL;
20.    if(predecessor != NULL)
21.          predecessor.lSn = root;
22.    CW(root.rSn.lSn, root.rSn);
23.    CW(root.lSn, root);
24.   }
25.   else
26.   {
27.    CW(root.lSn, root);
28.    define rightmost = root.lSn;
29.    while(rightmost.rSn != NULL)
30.          rightmost = rightmost.rSn;
31.    rightmost.rSn = root;
32.    define newroot = root.lSn;
33.    root.lSn = NULL;
34.    root = newroot;
35.    if(predecessor != NULL)
36.          predecessor.lSn = root;
37.    CW(rightmost.rSn, rightmost);
38.   }
39.  }
40.  }
41. }
```

b)

***Figure 2.*** *The algorithms for a) CCW() and b) CW()[5]*

plexity analysis will be outlined, resulting in the worst- and best-case scenarios for the algorithm, stated in the form of presenting them as functions of the number of nodes in the tree *n*. Afterwards, it will be shown how those results are tested empirically, addressing the analytical results for the space complexities of the worst case, best case and average case of the algorithm. The paper will end with a conclusion about the material presented herein.

**Space Complexity – Analytical Approach**

The analysis will be done upon the CCW() version of the algorithm, i.e. it will concern Figure 2a. As stated in [1], the CW() algorithm is an inverse of CCW()— substituting "left" for "right" and vice versa, as well as CCW() for CW() (for the recursive calls) will transform the CCW() algorithm into the CW() algorithm, so the following analysis can thus be used for the CW() algorithm as well. The number of nodes of the tree will be denoted by $n$. The line numbers will refer to the algorithm in Figure 2.

The space complexity is concerned with the memory which gets occupied during the execution of the algorithm. The occupied memory consists of the memory of the system stack given to every recursive call and to auxiliary variables which are needed for certain cases of the algorithm.

Let $S(n)$ denote the space complexity function of the binary tree roll. The trivial and the non-trivial cases of the CCW() algorithm will be analyzed for space complexity, since they all cause that a memory stack frame be used and placed on the call stack. In this analysis, the size of the stack frame reserved by the (recursive) function calls will be denoted by $s$ ($s > 0$), whereas the size of the auxiliary variables introduced by a certain invocation case (more specifically, the third case of the algorithm, specified by lines 25-38 in Figure 2) will be denoted as $a$ ($a > 0$).

The space complexity analysis cannot be performed using simple addition of function calls on the call stack, since, during the execution of the algorithm, some function calls get completed and leave the call stack, whereas others take their place. Therefore, it is the *depth* of *active function calls on the call stack* which is the measure of the space complexity [11], and a way to determine it needs to be devised. An example of such an approach is by drawing a *call stack tree* of the recursive function calls placed on the stack, which will be used in this paper. In the following analysis, such trees will be displayed for both the trivial and the non-trivial cases of the algorithm.

As will be shown in the following sections, the topology of the tree is a determining factor in the space complexity analysis of the algorithm. Since any topology of a tree with a given number of nodes $n$ is equally likely to be passed to the roll algorithm, it is unfeasible to derive an equation $S(n)$ as a function solely of the number of nodes $n$. However, it is possible to derive the worst- and best-case scenarios for the space complexity which are dependent on $n$ and such approach will be shown for the space complexity analysis. To do so, the concept of a *terminal situation* will be introduced, which is a situation in which a case of the algorithm is invoked after which there are no more recursive calls, except to the trivial case only. This approach will be used to determine the space efficiency of the binary tree roll algorithm and thus express the space complexity as a function of the number of nodes in the binary tree.

*The trivial case - line 3*

The trivial case gets invoked every time an empty tree is given to the CCW() algorithm for processing, i.e., when the test in line 3 of Figure 1 yields false. This case simply generates a function call in the call stack, which momentarily occupies space $s$ and then leaves the stack. It can be displayed as in Figure 3.

Since the CCW(0) case equals to a constant space complexity, its conversion to the constant $s$ will not be displayed in subsequent figures and will be assumed to happen instantaneously. Therefore, the trivial case will be assumed to be a part of the terminal situations of the other, non-trivial, cases.

$$CCW(0)$$
$$\downarrow$$
$$s$$

***Figure 3.*** *A call stack tree for the trivial case*

*First case - lines 5-10*

This case is invoked when the root has no right sub-tree (line 5 in Figure 2). The root's left sub-tree will be placed as its right sub-tree and a recursive call will be made upon the new right sub-tree. Figure 4 presents this case visually.

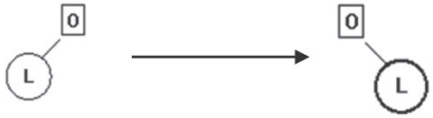



***Figure 4.*** *The first basic case in the CCW() algorithm [1]*

In the first case, there are no auxiliary variables; just a function call to the sub-tree which is rolled from left to right (as shown in Figure 4). The sub-tree has one node less than the initial tree, so this can be represented as in Figure 5.

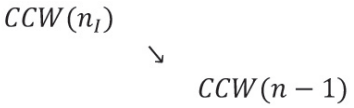



***Figure 5.*** *A call stack tree for the first case*

Note that the function call was denoted as $CCW(n_1)$. This is to stress that it is not only the amount of nodes in the tree that makes the difference, but also its *topology*. In other words, the first case will be invoked upon a tree with $n$ nodes (which may also be a sub-tree of the original tree) only if its topology is such that the root of that tree does not have a right sub-tree. If the topology is such that the root has a right sub-tree having no right sub-tree of its own, the second case will be invoked, whereas if the root has a right sub-tree having a right sub-tree of its own, the third case will be invoked.

A terminal situation in the first case happens when CCW() is invoked on a (sub-)tree with just a single node (i.e., a leaf of the tree). This situation shows when the first case will no longer be invoked and it can be represented as in Figure 6.

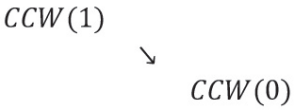



***Figure 6.*** *A call stack tree for the terminal situation of the first case (a tree with just one node)*

*Second case - lines 11-24*

This case is activated when the root of the tree has a right sub-tree, which in turn does not have a right sub-tree of its own (line 13 in Figure 2). It is shown visually in Figure 7.

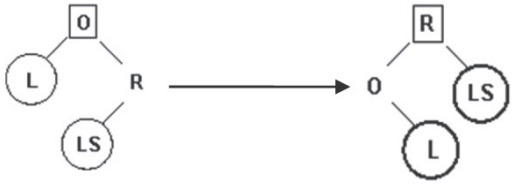



***Figure 7.*** *The second basic case in the CCW() algorithm [1]*

This case of the algorithm also does not introduce auxiliary variables, whereas it produces two recursive function calls on sub-trees which have a total of nodes. In other words, two of the nodes get handled by the second case, whereas the remainder of the nodes is processed by the subsequent recursive calls. This can be represented as in Figure 8:

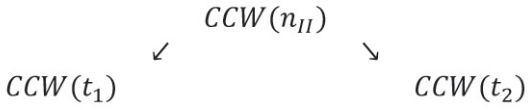



***Figure 8.*** *A call stack tree for the second case*

where $t_1 + t_2 = n - 2$.

A terminal situation in the second case happens when CCW() is invoked on a (sub-)tree containing two nodes, arranged as a topology of a root and its right child node. In such a situation, both of the nodes are handled by the second case and the two recursive calls are made on empty trees. This can be shown as in Figure 9.

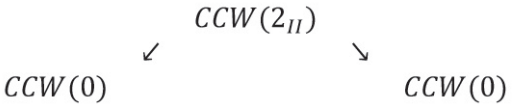



***Figure 9.*** *A call stack tree for the terminal situation of the second case (a tree containing just a root and its right child node)*

*Third case - lines 25-38*

This case gets invoked when there is a stem of two or more right child nodes to the root (i.e., when the root has a right sub-tree, which has a right sub-tree of its own and so on; line 25 in Figure 2). As stated in [1], this case deals with the downshift of stems of right child nodes and transforming them into stems of left child nodes, according to the principle of CCW binary tree roll. The algorithm first creates a recursive call upon the right child node of the root and it continues to do so until a basic case is reached (i.e., until a sub-tree with at most one right child node is reached, following the stem of right child nodes from the root towards its rightmost child node). When such case is handled by the algorithm, the remainder of the third case relocates the former root of the tree to be the leftmost child node in the newly rolled tree, and the procedure is then recursively invoked again on the former root (and its entire left sub-tree), now placed as the leftmost node in the sub-tree handled by the third case. Figures 10 and 11 show the third case visually.


***Figure 10.*** *The third and most complex case in the CCW() algorithm [1]*

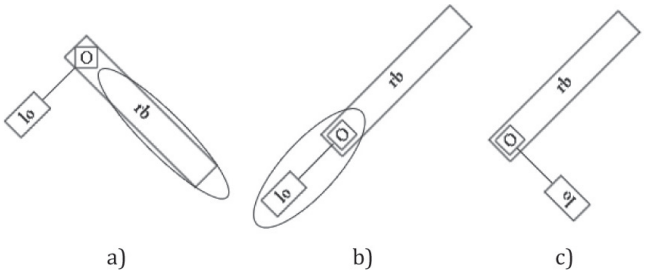



***Figure 11.*** *The third case of the CCW() algorithm: a) the head recursion (ellipse) of the third case deals with the stem of right child nodes () and transforms it into a stem of left child nodes via downshift; b) the root () is linked as the leftmost in the stem of left child nodes and the tail recursion (ellipse) of the third case is invoked upon it; c) since the former root does not have a right child node of its own, the tail recursion will invoke the first case, and the left sub-tree of the former root () will become its right sub-tree*

Concerning the space complexity, the third case of the algorithm includes a head recursion on the nodes not containing the root and its left sub-tree, two auxiliary variables and a tail recursion on the root and its left sub-tree. This can be displayed as in Figure 12, from where it can be seen that the third case does not perform the actual roll, but only the downshift of the tree (since the amount of processed nodes does not decrease in the subsequent recursive calls).

$$\begin{array}{ccc} & CCW(n_{III}) & \\ \swarrow & & \searrow \\ CCW(n-1-lo) & & 2a+CCW(1+lo) \end{array}$$

***Figure 12.*** *A call stack tree for the third case;  is the number of nodes in the left sub-tree of the root*

A terminal situation for the third case is when CCW() is invoked on a (sub-)tree with three nodes which form a stem of right child nodes. In such a situation, the head recursion is invoked upon the bottom two nodes, which are handled by the second case, and the tail recursion is invoked upon the root, which is handled by the first case. This is graphically represented in Figure 13.

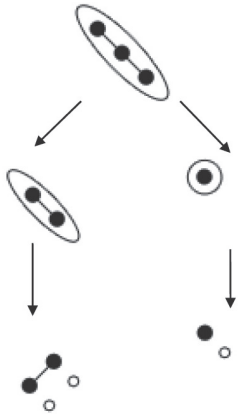


***Figure 13.*** *A terminal situation for the third case. The head recursion is invoked on the bottom two nodes of the stem (left-hand side) which get CCW rolled and induce two recursive calls on empty sub-trees (small circles to the lower right of each node). The tail recursion gets invoked on the root of the stem, after the downshift process (right-hand side), which finishes with a recursive call on an empty sub-tree (small circle)*

The call stack tree of the terminal situation can therefore be shown as in Figure 14.

$$\begin{array}{ccccc} & & CCW(3_{III}) & & \\ & \swarrow & & \searrow & \\ & CCW(2_{II}) & & 2a+CCW(1) & \\ \swarrow & & \searrow & & \searrow \\ CCW(0) & & CCW(0) & & CCW(0) \end{array}$$

***Figure 14.*** *A call stack tree for the terminal situation of the third case*

*The space complexity analysis: the space efficiency parameter*

As initially assumed in the space complexity analysis, every recursive call ($CCW(n_I)$), ($CCW(n_{II})$), ($CCW(n_{III})$), ($CCW(n-1)$ etc.) occupies the same amount of memory space. The maximum depth of the call stack tree for a given roll operation therefore shows the maximum amount of memory space occupied by the recursive calls during the algorithm execution, i.e., represents the upper bound of the space complexity of the algorithm for a given tree with  nodes. This can be used to determine the extreme scenarios, i.e., the worst- and best-case scenarios, for the space complexity of the CCW() algorithm.

To better quantify this, it is useful to introduce the *space efficiency* of the algorithm, as in Equation 3:

$$se_{case} = \frac{number\ of\ nodes\ processed\ per\ terminal\ situation\ of\ a\ given\ case}{amount\ of\ memory\ occupied\ by\ the\ terminal\ situation\ of\ the\ given\ case} \quad (3)$$

As can be inferred from Figures 9, 12 and 14, $se_I = \frac{1}{2s}$ (in the first case, processing 1 node takes 2 frames of recursive calls), $se_{II} = \frac{2}{2s} = \frac{1}{s}$ (in the second case, processing 2 nodes takes 2 frames of recursive calls), and $se_{III} = \frac{3}{3s+2a}$ (in the third case, processing 3 nodes takes 3 frames of recursive calls, plus the memory occupied by the auxiliary variables), respectively. Since every recursive call allocates space for two formal parameters (root and predecessor) of the same type as the auxiliary variables introduced in the third case of the algorithm (leftmost and newroot in lines 28 and 32 of Figure 1, respectively), it can be stated that the amount of space occupied by the auxiliary variables is, at best, equal to the space occupied by the recursive call, i.e., $2a \leq s$. This yields that $\frac{1}{s} = \frac{3}{3s} > se_{III} = \frac{3}{3s+2a} \geq \frac{3}{4s}$, and, since $\frac{1}{s} > \frac{3}{4s} > \frac{1}{2s}$, it follows that $se_{II} > se_{III} > se_I$. In other words, as far as space complexity of the algorithm is concerned, the most efficient is the second case and the least efficient is the first case. Therefore, the best-case scenario happens when a tree is processed solely by the second case of the algorithm and the worst-case scenario happens when a tree is processed solely by the first case of the algorithm. Figure 15 shows trees with $n = 6$ which will exhibit worst-case and best-case space complexities when rolled by the CCW() algorithm.

a) b)

***Figure 15.*** *Trees with which, when rolled with the CCW() algorithm, will exhibit a) worst-case time complexity; and b) best-case time complexity*

In the worst case, there will be as many recursive calls as there are nodes in the tree, plus one for the trivial case, which means that the equation for the worst-case scenario is represented by Equation 4:

$$S_O(n) = n + 1 \quad (4)$$

It is straightforward to notice that this equation has a linear time complexity, which means that the worst-case scenario for the space complexity of the algorithm is linear, or, as stated in Equation 5:

$$S(n) = O(n) \quad (5)$$

To analyze the best case, it is important to understand the regularity by which a tree is formed, which will be processed solely by the second case of the algorithm. The smallest such tree has $n = 2$ and consists of a root and its right child node. The next "best-case tree" will be formed by connecting two such sub-trees, consisting of a root and its right child node, as left child nodes of both the original root and the right child node of the root, which yields a tree with $n = 6$ (as shown in Figure 15b). The next such "best-case tree" will be formed by connecting sub-trees consisting of a root and its right child node to all nodes not having left child nodes, which are four in the "best-case tree" for $n = 6$, which will result in a new "best-case tree" with $n = 14$ (as shown in Figure 16). In such tree, there are eight nodes not having left child nodes, so connecting sub-trees consisting of a root and its right child node to all eight nodes will yield a new "best-case tree" with $n = 30$ and so on.

***Figure 16.*** *Tree with which, when rolled with the CCW() algorithm, will exhibit the best-case time complexity. It is obtained by linking every node, which does not have a left sub-node, from the previous such tree (shown in Figure 15b), with a sub-tree consisting of a root and its right sub-node. Such sub-trees are indicated by ellipses around them*

From the following, it can be concluded that the initial "best-case tree" has $n_0 = 2$ nodes, and each successive "best-case tree" has $n = \sum_{i=1}^{j} 2^i = 2^{j+2} - 2$ nodes, where $j$ is the length of the stem of left sub-nodes starting from the root (or from its right child node), which is the longest such stem (note that this stem does *not* include the root). In that case, $j + 2$ is the number of stack frames used for the recursive calls: $j$ for the calls upon sub-trees with roots in the stem (there are $j$ such nodes in the stem, excluding the root), 1 for the call on the initial root of the tree and 1 for the final call, on an empty tree from the leftmost node in the stem. Thus, the equation for it represents the relation for the best-case space complexity, expressed in terms of $n$, and it can be stated as in Equation 6:

$$S_{\Omega}(n) = s \cdot \lg(n + 2) \tag{6}$$

where $s$ is the amount of memory used by every invocation of the recursive function in the system stack. Equation 6 indicates that the best-case space complexity is tightly logarithmic. This can be stated as in Equation 7:

$$S_{\Omega}(n) = \Theta(\lg n) \tag{7}$$

An equivalent statement is that the space complexity of the CCW() algorithm is logarithmic in the best case, as stated in Equation 8:

$$S(n) = \Omega(\lg n) \tag{8}$$

It is thus shown that the space complexity of the binary tree roll algorithm has a linear worst-case space complexity (Equation 5) and a logarithmic best-case space complexity (Equation 8).

**Space Complexity – Empirical Approach**

In order to be certain about how much space is needed to perform CCW roll on a tree with nodes, an exhaustive analysis needs to be performed. This includes obtaining all topologies of binary trees with nodes and performing CCW roll on them, while obtaining the stack depth for each tree while it is being CCW rolled. For this, it is necessary to first generate all topologies of binary trees for a given and then execute CCW roll on all of them, while measuring the stack depths during the executions of the CCW roll. The smallest value of the stack depth while CCW rolling a tree with nodes will represent the best case for that , whereas the largest value of such a stack depth will represent the worst case for that . As following from the theoretical analysis, the best cases for increasing values of are expected to grow logarithmically, and the worst cases are expected to grow linearly. It is also appealing to know whether the space complexity of the algorithm would be more dominantly logarithmic or linear, which is why an average time complexity would also need to be extracted, as an average of the space complexities for all topologies of binary trees for a given number of nodes.

In order to obtain all topologies of binary trees with a given number of nodes, the Catalan Cipher Vector approach is used in this paper. A Catalan Cipher Vector [6] is a vector which uniquely determines

a binary tree's topology. For a tree with  nodes, there will be  (n-th Catalan number) topologies of binary trees and thus  Catalan Cipher Vectors. Table 1 shows all the ranks, their corresponding Catalan Cipher Vectors, and the appropriate binary trees, for n = 4 nodes.

***Table 1.*** *Ranks and enumerations of the binary trees with  nodes using the Catalan Cipher Vector approach*

| Rank | Catalan Cipher Vector | Binary Tree |
|---|---|---|
| 0 | [0 1 2 3] |  |
| 1 | [0 1 2 4] |  |
| 2 | [0 1 2 5] |  |
| 3 | [0 1 2 6] |  |
| 4 | [0 1 3 4] |  |
| 5 | [0 1 3 5] |  |
| 6 | [0 1 3 6] |  |
| 7 | [0 1 4 5] |  |
| 8 | [0 1 4 6] |  |
| 9 | [0 2 3 4] |  |
| 10 | [0 2 3 5] |  |
| 11 | [0 2 3 6] |  |
| 12 | [0 2 4 5] |  |
| 13 | [0 2 4 6] |  |

Since the initial Catalan Cipher Vector for a tree with  nodes is always  [6], it is possible to generate the corresponding binary tree for it, and determine the stack depth occupied during the execution of CCW() on it. Then, the subsequent Catalan Cipher Vector can be obtained, the corresponding binary tree can be generated from it, have CCW() executed on it and determine the stack depth needed and so on, until all  binary tree topologies get processed this way. The obtained minimum depth represents the best case, the maximum depth represents the worst case, whereas the average case is calculated as the sum of all depths divided by the number of possible trees with  nodes.

The results for such an analysis have been performed and the results are given in Table 2.

***Table 2.*** *Stack depths necessary to perform CCW() on all topologies of binary trees with given numbers of nodes*

| $n$ | $C(n)$ | *Min* | *Max* | *Avg* | *Total* |
|---|---|---|---|---|---|
| 2 | 2 | 2 | 3 | 3 | 5 |
| 3 | 5 | 3 | 4 | 3 | 16 |
| 4 | 14 | 3 | 5 | 4 | 54 |
| 5 | 42 | 3 | 6 | 4 | 187 |
| 6 | 132 | 3 | 7 | 5 | 664 |
| 7 | 429 | 4 | 8 | 6 | 2.393 |
| 8 | 1.430 | 4 | 9 | 6 | 8.719 |
| 9 | 4.862 | 4 | 10 | 7 | 32.073 |
| 10 | 16.796 | 4 | 11 | 7 | 118.848 |
| 11 | 58.786 | 4 | 12 | 8 | 443.081 |
| 12 | 208.012 | 4 | 13 | 8 | 1.660.503 |
| 13 | 742.900 | 4 | 14 | 8 | 6.250.670 |
| 14 | 2.674.440 | 4 | 15 | 9 | 23.620.379 |
| 15 | 9.694.845 | 5 | 16 | 9 | 89.560.477 |
| 16 | 35.357.670 | 5 | 17 | 10 | 340.599.877 |
| 17 | 129.644.790 | 5 | 18 | 10 | 1.298.763.168 |
| 18 | 477.638.700 | 5 | 19 | 10 | 4.964.255.082 |

The results are interpreted as follows. In the first data row, for a tree with  = 2 nodes (first column), there are  = 2 (second column) total topologies of binary trees. Executing CCW() on all of them yields a  (sixth, i.e. last column) of 5 levels of stack depth, leading to an  (average – fifth column) of 3 levels of

stack depth per binary tree topology. Of all topologies, the (minimum – third column) levels of stack depth necessary to complete CCW() on a binary tree topology with 2 nodes is 2, and the (maximum – fourth column) number of such levels is 3. This interpretation follows all rows of the table, up to and including binary tree topologies for = 18 nodes.

Plotting the results obtained from Table 3 results in a chart like in Figure 17.

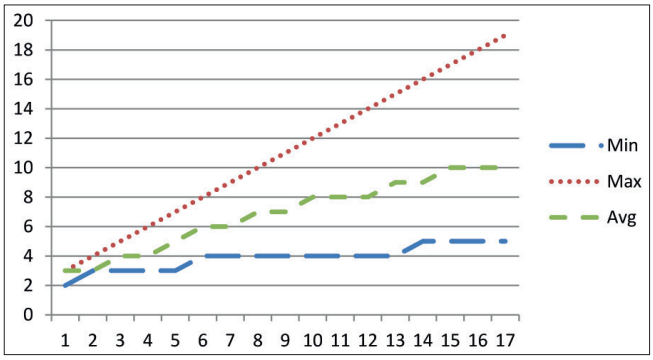



***Figure 17.*** *A plot of the results given in Table 3*

Interpolating the equations for the worst and best cases yields the equations and respectively, confirming the conclusions obtained in the analysis of the space complexity that the worst-case space complexity of the CCW() algorithm is linear, and its best-case space complexity is logarithmic. Interpolating for the average case results in the equation , meaning that the average space complexity of the CCW() roll algorithm is dominantly logarithmic. Therefore, it can be stated that, on average, the space complexity of the CCW() algorithm is logarithmic.

## Conclusion

This paper presented an analysis of the space complexity of the binary tree roll algorithm, specifically its counterclockwise (CCW()) variant, with the note that the analysis for its clockwise (CW()) variant is analogous. For it, the theoretical analysis was derived in such way as to show the amount of memory occupied during the execution of the algorithm, shown primarily through the number of frames occupied in the call stack by the recursive function calls. The analysis of the space efficiency of each of the three cases of the algorithm, as well as the trivial case, showed that trees which are processed solely by the first case of the algorithm (invoked upon a node which has no right sub-tree) yield the worst space complexity, whereas trees which are processed solely by the second case of the algorithm (invoked upon a node which has a right sub-tree, which in turn has no right sub-tree of its own) yield the best space complexity. The equations were derived to be linear for the worst case and logarithmic for the best case. An exhaustive empirical analysis was performed, counting the minimum and maximum stack depths for all trees with given numbers of nodes . The equations in the theoretical analysis were confirmed by the empirical results, that the CCW() algorithm has linear space complexity for the worst case and logarithmic space complexity for the best case. Moreover, the average case analysis showed that the CCW() algorithm has a dominantly logarithmic space complexity, so it can be stated that its average space complexity is logarithmic.

## References

[1] Adel'son-Vel'skii G. M. and Landis E. M. (1962) "An algorithm for the organization of information," *Soviet Math. Doklady*, 3: 1259-1263

[2] Amani M., Lai K. A. and Tarjan R. E. (2016) "Amortized rotation cost in AVL trees," *Information Processing Letters*, 98(5): 327-330

[3] Bayer R. and McCreight E. (1972) "Organization and Maintenance or Large Oriented Indexes," *Acta Informatica*, 3(3): 173-189

[4] Bozinovski A. and Bozinovski S. (2004) N-queens Pattern Generation: An Insight Into Space Complexity of a Backtracking Algorithm, Proceedings of the Third International Symposium on Information and Communication Technologies, Las Vegas, NV, USA (pp. 281-286). ACM

[5] Božinovski, A. and Ackovska, N. (2012) The Binary Tree Roll Operation: Definition, Explanation and Algorithm, International Journal of Computer Applications, 46(8):40-47

[6] Božinovski, A., Stojčevska, B. and Pačovski, V. (2013) Enumeration, Ranking and Generation of Binary Trees Based on Level-Order Traversal Using Catalan Cipher Vectors, Journal of Information Technology and Applications, 3(2):78-86

[7] Božinovski, A., Tanev, G., Stojčevska, B., Pačovski, V. and Ackovska, N. (2016) Time Complexity Analysis of the Binary Tree Roll Algorithm, International Journal of Computer Applications, 6(2):53-62

[8] Brassard G. and Bratley P. (2002) *Fundamentals of Algorithmics*. Prentice Hall of India

[9] Buchmann J., Dahmen E., Klintsevich E., Okeya K. and Vuillaume C. (2007) "Merkle signatures with virtually unlimited signature capacity," In *Proc. 5th Int. Conf. Applied Cryptography and Network Security*, Zhuhai, China (pp. 31-45). LCNS Springer

[10] Cormen T. H., Leiserson C. E., Rivest R. L. and Stein C. (2009) *Introduction to Algorithms, 3rd ed.*, The MIT Press

[11] Goodrich, M. T. and Tamassia, R. (2008) *Data Structures and Algorithms in Java, Third Edition*, Wiley India Pvt.

[12] Katz J. (2003) "Binary Tree Encryption: Constructions and Applications," In *Information Security and Cryptology (ICISC 2003)*, 2971: 1-11, LCNS Springer

[13] Kreher D. L. and Stinson D. R. (1998) *Combinatorial Algorithms: Generation, Enumeration, and Search*. Discrete Mathematics and its Applications (Book 7), CRC Press, 1st Edition

[14] Lee Y. and Lee J. (2015) "Binary tree optimization using genetic algorithm for multiclass support vector machine," *Expert Systems with Applications*, 42(8): 3843-3851

[15] Li Y., Xu M., Zhao H. and Huang W. (2016) "Hierarchical fuzzy entropy and improved support vector machine based binary tree approach for rolling bearing fault diagnosis," *Mechanism and Machine Theory*, 98: 114-132

[16] Liu B., Shen Y., Chen X., Chen Y. and Wang X. (2014) "A partial binary tree DEA-DA cyclic classification model for decision makers in complex multi-attribute large-group interval-valued intuitionistic fuzzy decision-making problems," *Information Fusion*, 18: 119-130

[17] Rich E. (1983) *Artificial Intelligence*. McGraw-Hill series in artificial intellilgence, McGraw-Hill Inc.

[18] Roch S. and Steel M. (2014) "Likelihood-based tree reconstruction on a concatenation of alignments can be statistically inconsistent, " *Theoretical Population Biology,* 100: 56-62

[19] Sedgewick R. (1998) *Algorithms in C++, Parts 1-4: Fundamentals, Data Structures, Sorting, Searching, 3rd ed,* Addison-Wesley

[20] Suh I. and Headrick T. C. (2010) "A comparative analysis of the bootstrap versus traditional statistical procedures applied to digital analysis based on Benford's Law," *Journal of Forensic and Investigative Accounting*, 2(2): 144-175.

Submitted: December 1, 2016.
Accepted: December 5, 2016.

## About the Authors

**Adrijan Bo**žinovski works as an Associate Professor at the School of Computer Science and Information Technology at University American College Skopje, where he is currently the Dean. He obtained his BSc from University "St. Cyril and Methodius" in Skopje, Macedonia, and his MSc and PhD from University of Zagreb, Croatia.

**Biljana Stojčevska** works as an Associate Professor at the UACS School of Computer Science and Information Technology. She received her BSc, MSc and PhD degrees in Computer Science at the Institute of Informatics, Faculty of Natural Sciences and Mathematics, at "Sts. Cyril and Methodius University" in Skopje, Macedonia.

**Veno Pachovski** (1965) graduated, completed MSc and got his PhD from Faculty of Natural Sciences and Mathematics, University "Sts. Cyril And Methodius", Skopje, Macedonia. Since 2009, he teaches a variety of courses at the University American College – Skopje, mainly within the School of Computer Sciences and Information technology (SCSIT).

**George Tanev** is an MSc graduate student of the School of Computer Science and Information Technology at University American College Skopje, Macedonia, where he acquired his BSc in Computer Science. Also works as a software developer in Skopje, Macedonia.

**Nevena Ackovska** is Associate Professor at the Faculty of Computer Science and Engineering at "St. Cyril and Methodius" University in Skopje, Macedonia. She holds B.Sc. in Computer Engineering, Informatics and Automation from Electrical Engineering Faculty (2000), M.Sc. in Bioinformatics (2003) and a Ph.D. in Bioinformatics (2008) from Faculty of Natural Sciences and Mathematics at "St. Cyril and Methodius University" in Skopje, Macedonia.

# THE USE OF DIGITAL SIGNATURE IN ELECTRONIC COMMUNICATION IN BIH - RESEARCH

**Tijana Talić**

*Paneuropean University APEIRON, Banja Luka, tijana.z.talic@apeiron-edu.eu*

**Contribution to the state of the art**

DOI: **10.7251/JIT1701020T** UDC: **004.056.55(497.6)**

**Abstract:** The increasing use of electronic mail for identity theft and unsolicited marketing and frequent presence of viruses as well, reduced the credibility of email as a communication tool. Authentication of the sender is well known defense against such attacks. One of the methods to ensure that authentication, secure communication via e-mail, is the use of digital signature.

**Keywords:** a digital signature, authentication, the education sector, business users.

## INTRODUCTION

The information society development in Bosnia and Herzegovina is lagging far behind, even in comparison with countries in the region, Europe and globally. One of the essential conditions for integration into the European and global business environment is the information society development in order to facilitate business and create conditions for faster economic growth of the country.

Efficient e-business is not possible if the security of electronically transmitted data is not ensured.

In this research we examined users' opinion about the importance and use of digital signatures in our environment. We split the users into two groups by their use of electronic communications and agencies in which they work:

1. Users employed in the education sector and
2. Business users

The first group consists of employees in the education sector. These are the users employed in educational institutions who know the subject matter, but the frequency of using electronic mail is relatively low. Generally, they can be considered as average users.

The second group consists of users to whom electronic communication is of crucial importance. The frequency of using electronic mail is high. In our case, this group consists mainly of civil servants and local government officials as well as of a number of business users who are related to accounting, transport and trade. This group is particularly vulnerable to the message authenticity.

The aim of the research was to examine users' attitudes toward:

- Security of e-mail communication
- Use of digital signatures and
- The significance of digital signatures in business communication.

The aim was also to determine if there were different attitudes toward the use and importance of digital signatures between the user employed in the stable business environment, including local government employees and those employed in educational institutions.

## RESEARCH RESULTS

The study included a total of 108 respondents aged between 21 and 65 years old.

| PROFESSIONAL QUALIFICATIONS | |
|---|---|
| Elementary school | 0 |
| Secondary school | 12 |
| College degree | 2 |
| University degree | 66 |
| Master or PhD | 28 |

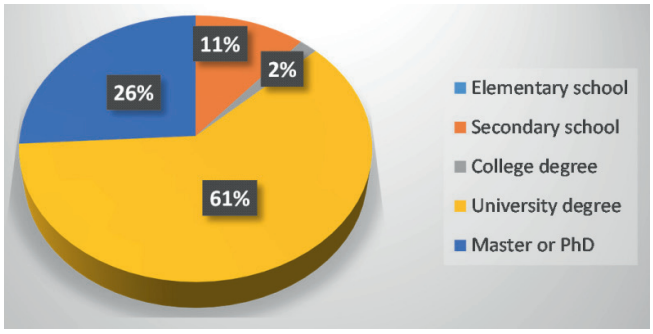



***Chart 1.***

Chart No 1 shows that the majority of respondents has a university degree, followed by Masters and PhDs.

The survey consisted of four groups of questions.

The first group of questions were answered by all respondents with the aim to determine the users attitudes on e-mail safety and reliability, as well as the use of digital signatures. We compared the answers given by the average users to the respondets that were considered as business users. Also, we investigated how many respondents used digital signature.

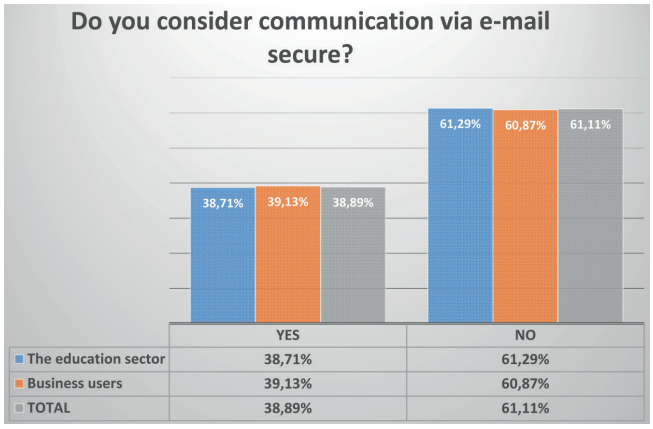



***Chart 2.***

The chart No. 2 shows that there are no large differences in the responses to the question, "Do you consider communication via e-mail secure?" given by business users in relation to the answers given by the teachers. In both groups, 60% of respondents believe that communication via e-mail is not secure.

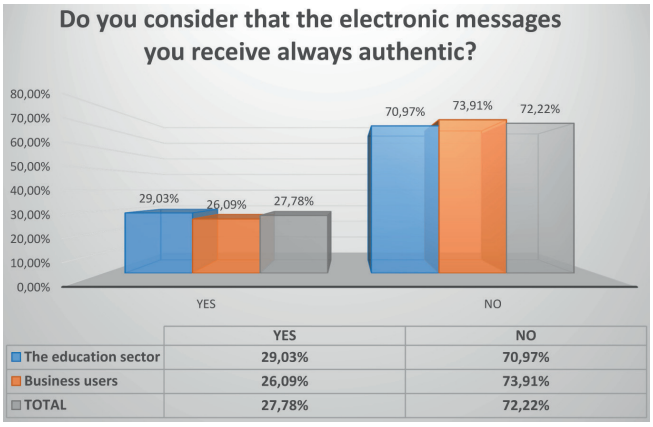



***Chart 3.***

The second question was about the authenticity of electronic messages. Answers to this question showed no big difference between the two groups of respondents as well.

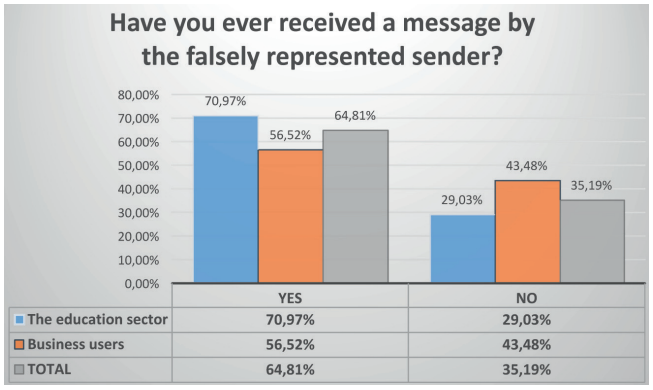



***Chart 4.***

Chart No 4 shows that 70.97% of education sector employees, and 56.52% of business users replied with '' Yes'' to the queston "Have you ever received a message by the falsely represented sender?"

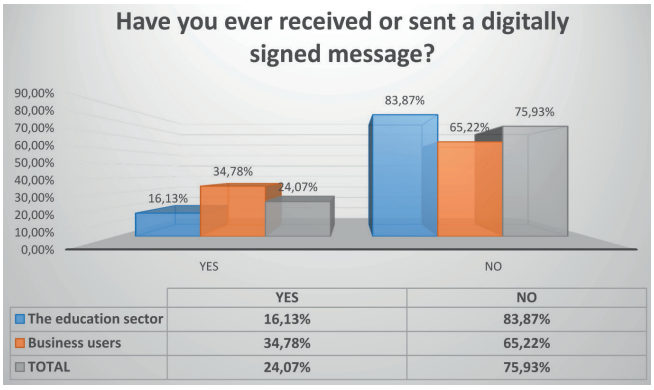
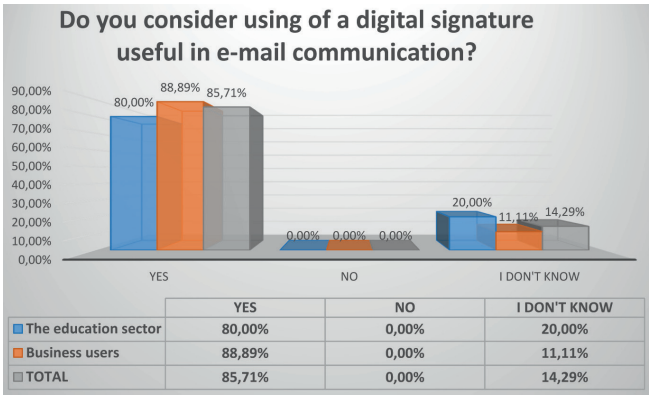



***Chart 5.***

Chart No 5 shows that only 16.13% of the education sector employees and 36.36% of business users, at least once, used the digital signature.

The second group of questions was answered only by the users who had, at least once, used a digital signature. The aim was to investigate the prevalence, importance and purpose of digital signature.



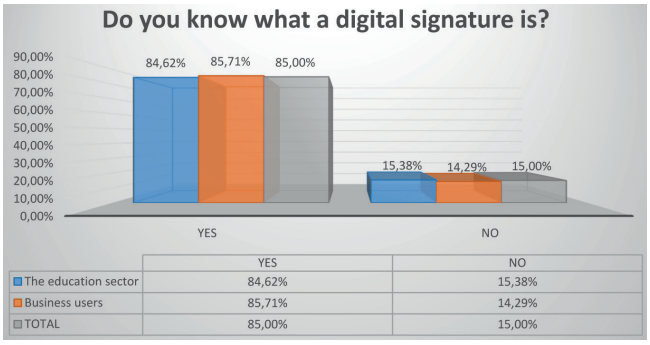

| | YES | NO | I DON'T KNOW |
|---|---|---|---|
| The education sector | 80,00% | 0,00% | 20,00% |
| Business users | 88,89% | 0,00% | 11,11% |
| TOTAL | 85,71% | 0,00% | 14,29% |

***Chart 6.***

The chart No. 6 shows that 85.71% of respondents who use digital signature consider it as a useful tool for e-mail communication.

Replies of the surveyed users of digital signature to the question "For what purpose a digital signature is used most often? " were as follows: e-mail signature, e-payment, signing documents and tax returns.

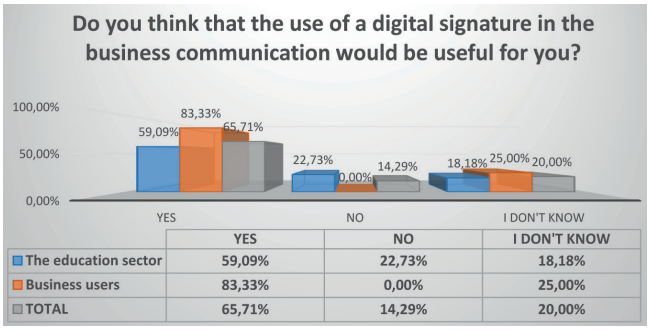
The third group of questions was only for those who had never sent nor received digitally signed message.




| | YES | NO |
|---|---|---|
| The education sector | 84,62% | 15,38% |
| Business users | 85,71% | 14,29% |
| TOTAL | 85,00% | 15,00% |

***Chart 7.***

The chart No. 7 shows that, according to the question "Do you know what a digital signature is?" there were no large differences between respondents who work in the education sector and the resondents who were considered as business users. 85% of the total number of respondents who had never used a digital signature knew what a digital signature was.

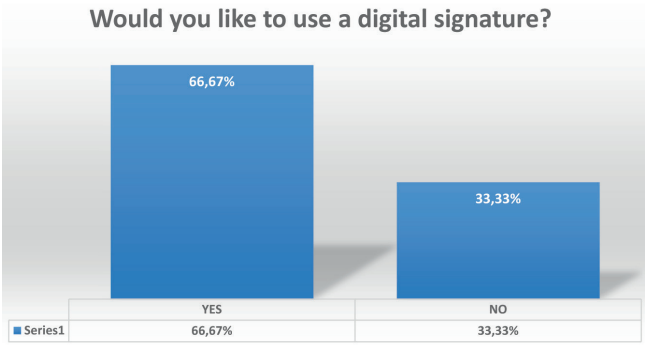



| | YES | NO | I DON'T KNOW |
|---|---|---|---|
| The education sector | 59,09% | 22,73% | 18,18% |
| Business users | 83,33% | 0,00% | 25,00% |
| TOTAL | 65,71% | 14,29% | 20,00% |

***Chart 8.***

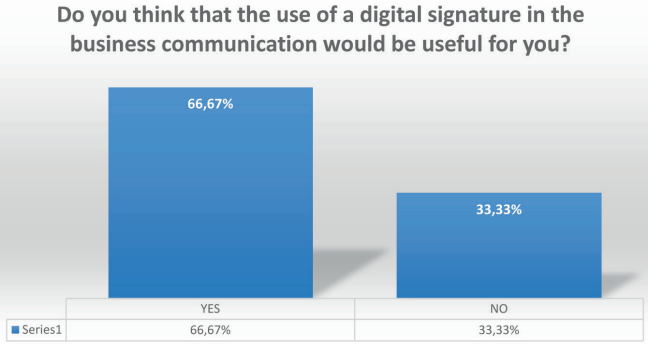
The chart No. 8 shows that in relation to the education sector employees, significantly more employees in the business sector believed that the use of digital signatures would be useful for them.

The fourth, and the last group of questions was answered only by respondents who answered the question "Do you know what a digital signature is?" with "NO". Before answering the questions, they had to read a short text in order to learn the meaning and use of digital signature.




| | YES | NO |
|---|---|---|
| Series1 | 66,67% | 33,33% |

***Chart 9.***




| | YES | NO |
|---|---|---|
| Series1 | 66,67% | 33,33% |

***Chart 10.***

The chart No. 9 and 10 show that 66.67% of respondents, who did not know what a digital signature was, wanted to use a digital signature and believed that its use would be conductive for them in business communication.

## Conclusion

This research determined the use of digital signature in BiH with special emphasis on Banja Luka area as high administrative and business center in BiH.

We also analyzed the results according to the activities of the respondents. We were particularly interested in two groups that were clearly split. Those are:

- Users employed in the education sector. These users can be considered as average users of electronic communication.
- Users from the business sector. We have grouped the staff members both in public administration and in trade, transport and services type of the accounting offices. In the second group the reliability and authentication of the sender is of crucial importance.

We analyzed the respondents' opinions about the importance of digital signature in electronic communication. The results show that the surveyed business users in our area do not consider that communication by e-mail is completely safe, nor the messages they receive are always credible.

Comparing to the employees in the education sector we note that a greater percentage of employees in the business sector are more likely to use a digital signature and they believe it is useful in business communication. The reason for this result is that commercial communication via electronic media, due to the type of work, requires frequent use of digital signature every day. On the other hand, jobs performed by employees in education sector, still do not require the use of digital signature, even though its use could indeed facilitate the administrative work in educational institutions.

Generally speaking, this study shows that the users are interested in applying the digital signature and the majority believes that its use would be beneficial for them as well.

This study shows that further research in this area of business communication is justified.

The aim of this research is to analyze and propose the best solution for authentication and applicaion of digital signature inside an educational institution.

## Literature

[1] Talic, T. (2016) Uloga i značaj digitalnog potpisa u poslovnoj komunikaciji, 8th International Conference „Information Technologies in e-Education"

[2] https://gnupg.org/gph/en/manual.html last accesed 23.05.2017.

Submitted: May 22, 2017.
Accepted: May 28, 2017.

## About the Author

**Tijana Talić** works as a Senior Assistant at Paneuropean University „Apeiron" and as a Professor of Mathematics and Computer Science in Secondary School Center „Gemit-Apeiron". She gratuated at The Faculty of Mathematics and Natural Science in 2008 and in 2014 she became Mr. Sci in Information Technology.

# EXPERT SYSTEMS IN A CLOUD COMPUTING ENVIRONMENT MODEL FOR FAST-PACED DECISION MAKING

**Mihalj Bakator, Dragica Radosav**
*University of Novi Sad, Technical faculty "Mihajlo Pupin, Zrenjanin*
*mihaljbakator@gmail.com, radoosav@tfzr.uns.ac.rs*

**Contribution to the state of the art**
DOI: **10.7251/JIT1701024B** UDC: **004.67/82:005.5**

**Abstract:** In this paper the use of cloud computing technologies and expert systems will be analyzed. Furthermore, the use of expert systems in a cloud computing environment will be addressed. Specifically a Cloud-Based Expert System (CBES) model for decision making will be presented. The mentioned model will include the model's infrastructure and its application. In addition, a theoretical approach will be used as a basis for the research and analysis. The CBES model offers effective, fast and reliable support for individuals or organizations when it comes to fast-paced decision making.

**Keywords:** cloud computing, environment, CBES model, decision, expert system.

## INTRODUCTION

In modern society, information technology has become a necessity. The Internet has revolutionized social interactions and the way of how people communicate. Furthermore, globalization and the age of digitalization have brought dynamic and fast information distribution. People can easily send, receive and share digital content. Cloud computing can be defined as a highly scalable virtual resource development model [18]. In addition, cloud computing technologies provide a wide set of options for flexible data storage. Nonetheless, cloud environments can "evolve" in various ways. Workflow scheduling research based on Modified Critical Path (MCP) and Dynamic Critical Path (DCP) algorithms can be used inside the cloud environment [9]. Likewise, expert systems can be implemented as well as other applications [16]. This kind of integration and use of expert systems depends on the structure and purpose of the cloud computing environment.

Expert systems are computer systems which emulate the decision making process of a human expert [17]. It has a very wide application in medicine, decision-making models, diagnostics and other types of services [21]. Expert systems can be integrated as a part of a cloud computing environment. Moreover, detailed and massive databases can enrich expert systems in the cloud environment. This can be achieved through real-time storage in the database with knowledge from experts. In addition, the added data is available to other users almost instantly [16]. Similarly, other services and applications can be provided alongside expert system support. Cloud computing environments are known for their flexibility and scalability [2]. Together, with expert systems, a model for fast-paced decision making can be created. In the modern business world, the dynamic markets require a fast decisions regarding product distribution, product manufacturing and achieving goals. Similarly, in medicine, in the automotive industry and manufacturing processes, effective di-

agnostics are important for good and fast decision making.

In this paper, a model of an expert system in a cloud computing environment for fast-paced decision making will be presented. The model includes the infrastructure of the cloud computing environment and the expert system model. Furthermore, details of the expert system, cloud platform and information flow patterns will be defined and presented. The paper addresses three main concepts. First, the principles and applications of expert systems will be defined. Second, the cloud computing environment will be presented. This includes the characteristics, architecture, advantages and disadvantages. The third section of the paper defines the elements of the model and shows the main concept of expert system implementation with cloud computing environment for fast-paced decision making.

This *Cloud-Based Expert System* model (CBES model) can improve the effectiveness of decisions and drastically reduces execution and decision making time.

Furthermore, the paper addresses the main concepts of cloud computing deploy models. In addition, software architectures and other types of cloud computing will be reviewed. The main goal of this paper is to present the possibilities of expert system use in a cloud computing environment. Furthermore, the benefits for small and medium enterprises (SME) are defined as well as the overall positive influence on society. The paper is motivated by the concept of decision making efficacy and the major role of cloud computing and expert systems in society. The presented CBES model offers an insight to the possibility of expert systems use in a cloud computing environment.

Expert systems and diagnostics

Expert systems are defined as tools which are capable to understand problems from a specific subject. The system uses acquired and available "knowledge" from databases to solve problems [12]. The "expertise" comes from experts who have knowledge about a specific subject. Furthermore, an expert system uses the experts' knowledge in order to solve a given problem [21]. Expert systems are based on two basic concepts. First, on an expert's knowledge which is not algorithmically organized, is stored into databases of the system. From here, the data is processed and used by the user through the user interface [11]. The second concept includes the expert's knowledge storing into the procedural system. This procedural knowledge can be modified and it generates new knowledge based on facts from the procedural system [11].

Expert systems can be presented as an intelligent "consultant" in a specific subject. The knowledge that an expert system contains is collected from one or more experts from one specific field. However, there are situations when expert systems can contain knowledge from two similar fields that together form a synergy [3] [11]. Known expert systems are:

- MYCIN, developed by Stanford University and is used for bacterial infection and antibiotics therapy diagnosis [19];
- MACSYMA, developed by Martin and Fateman and it is used for symbolic solving of differential and integral calculations [6];
- EXPERT, developed by Weiss, Kulikowski and Safir, and it is used for glaucoma diagnosis [4];
- CADUCES, developed by Pople, Myers and Miller, and it is used for internal statistics [14];
- GRAPH developed by Cvetković and others and it is used for medical diagnostics [7].

Expert system development includes one or more experts from whom the knowledge is collected. Problems may occur when the experts cannot present their knowledge in an adequate way. When developing an expert system, there are basic steps which are used for development. These steps are [11]:

- conceptual design;
- collecting knowledge;
- system development;
- estimating performance;
- estimating acceptability;
- demonstration of the prototype;
- revision and further development;
- expert system realization.

The main components of an expert system are the knowledge base, working base or database, inference mechanism and connection with the user [11]. The knowledge base includes knowledge from experts and this type of knowledge is presented as declarative knowledge [11]. The working database contains "initial" data and "flowing" data as well as facts which are connected to objects from the prob-

lematic field. The inference mechanism contains two main components. The first one is the execution mechanism and the second one is the control mechanism [15]. Without the inference mechanism the expert systems would be nothing more than a standard database. In addition, all knowledge from the database is processed through the inference mechanism [11] [15]. The simple scheme of an expert system is shown in Figure 1.

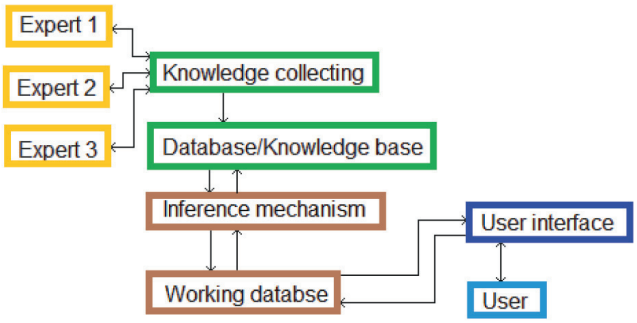



***Figure 1.*** *Expert system scheme (Source: own)*

One expert or more can store knowledge into the database for knowledge collecting. Next, the data is stored into the database and it is processed within the inference mechanism. Through the working base and user interface the user can access the processed data/knowledge.

Expert system can be used for diagnostics, medicine, debugging, monitoring, decision making and other. In this paper a decision making model will be presented in a cloud computing environment. The goal is to show the possibility of effective fast-paced decision making in a dynamic working environment.

The advantages of expert systems are ease of maintenance and rapid prototyping. The disadvantages of expert systems manifest as knowledge acquisition problems, performance when using tools as Lips and database integration for early expert systems [11].

In the next section, cloud computing technologies and cloud computing environment will be addressed. Further, the CBES model will be presented.

Cloud computing environment

Cloud computing can be defined as Interned based computing. This type of computing shares processing resources and data to computers or other devices. Furthermore, this kind of sharing is on demand. The main concept is based on the Internet, and the majority of the connected devices are through the Internet, thus making it very flexible and convenient [20]. The main web services delivered from the cloud are [16]:

- Communication-as-a-Service (CaaS)
- Infrastructure-as-a-Service (IaaS)
- Monitoring-as-a-Service (MaaS)
- Platform-as-a-Service (PaaS)
- Software-as-a-Service (SaaS)

Communication as a Service (CaaS) can be defined as an outsourced communications solution [20]. CaaS providers manage necessary hardware and software which are required for delivering VoIP services, Instant Messaging, Video calls and other types of communications [16] [20]. CaaS platforms are flexible and scalable. This is important for big organizations that need a manageable growth in their IT sector. Providers can usually handle peak loads when it comes to providing resources [22]. The main advantages of CaaS are shared costs with other vendors, remote management of the infrastructure and all-around communication support [16].

Infrastructure as a Service (IaaS) usually provides a virtualization platform as a service. The necessary computer hardware and software for such platforms are routers, firewalls, internet connectivity and platform virtualization machines. This type of equipment is expensive, thus providers sell the resources as virtual platforms for individuals or organizations [16] [20]. The main benefits of IaaS are ready access to preconfigured environment which is based on Information Technology Infrastructure Library (ITIL), has the latest technology equipment, secured computing platforms, manageable service-demand peaks, valleys and lower costs [16].

Monitoring as a Service (MaaS) provides security for business, enterprise and even government platforms. MaaS offers protection and security against internal and external threats [16].

Platform as a Service (PaaS) provides dynamic, secure and broad platforms for applications, downloads, developers, IT managers, and others. Businesses can focus on innovation, while the PaaS offers the infrastructure [16]. In addition, PaaS includes security options for various sectors in an organization.

Software as a Service (SaaS) includes applications that are ran on virtual platforms, from where

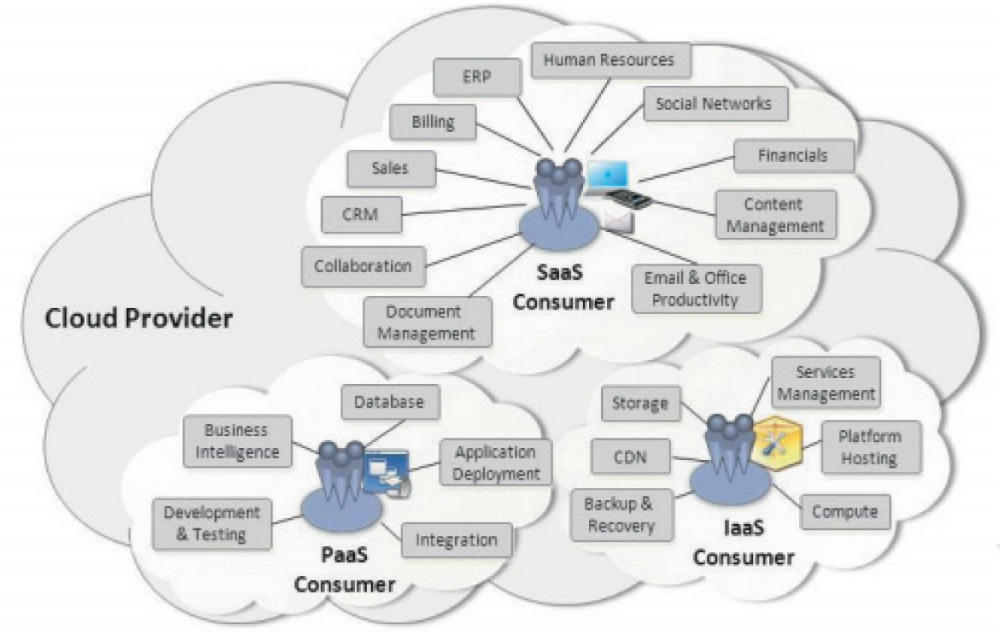



***Figure 2.*** *Options that SaaS, IaaS and PaaS offer to the consumers (Source: Bahrami & Singhal, 2015).*

users can access the functions of those applications [16] [20]. Therefore, SaaS provides a secure platform where resource "hungry" software can run "smoothly" and process the input which has been entered by the users.

The above mentioned cloud web services are usually brought down to IaaS, PaaS and SaaS. Nonetheless, the MaaS and CaaS are sometimes defined as a part of the SaaS or IaaS. Figure 2. shows the options that a SaaS, PaaS and IaaS consumers have in the cloud [1].

In Figure 2., it can be seen that PaaS provides business intelligence, databases, development and testing, integration and application deployment to the consumer. IaaS provides storage, SDN, backup and recovery, services management, platform hosting and computing resources for the consumer. Third, the SaaS provides ERP, Billing options, human resource management, social networks, financial report control, content management, email and office productivity, document management, collaboration CRM and sales options to the consumer.

Characteristics of cloud computing are improved agility of organizations, cost reductions by providers, infrastructure is provided by a third party, location independence, wide-range compatibility of devices, easy maintenance of applications, monitored performance, increased productivity, improved reliability, scalability and elasticity and improved security [17].

The main advantages of cloud computing environments are flexibility, disaster recovery, automatic software updates, free capital-expenditure, increased collaboration, work from anywhere, document control, security, competitiveness and environmentally friendly [16] [17] [18] [19]. Flexibility is manifested as the possibility of growth within the cloud platform without fluctuating bandwidth and resources. In addition, when businesses want to add new users and use more resources, there are little or no "freeze" loads at peak points.

Disaster recovery in a cloud environment is cost effective and lets small and medium enterprises (SME) secure their valuable data in case of data loss, fire or power surge. Furthermore, files and applications are backed up and secured from malicious attacks or other types of attacks. Automatic software updates are a big deal for SMEs, because a lot of them do not have the financial resources to hire IT administrators. Suppliers can take care of software updates remotely without the need to be physically present in the organization where the computers and devices are located.

Capital-expenditures are much lower compared to traditional servers. Companies do not need to buy high-end computers. Cloud computing environ-

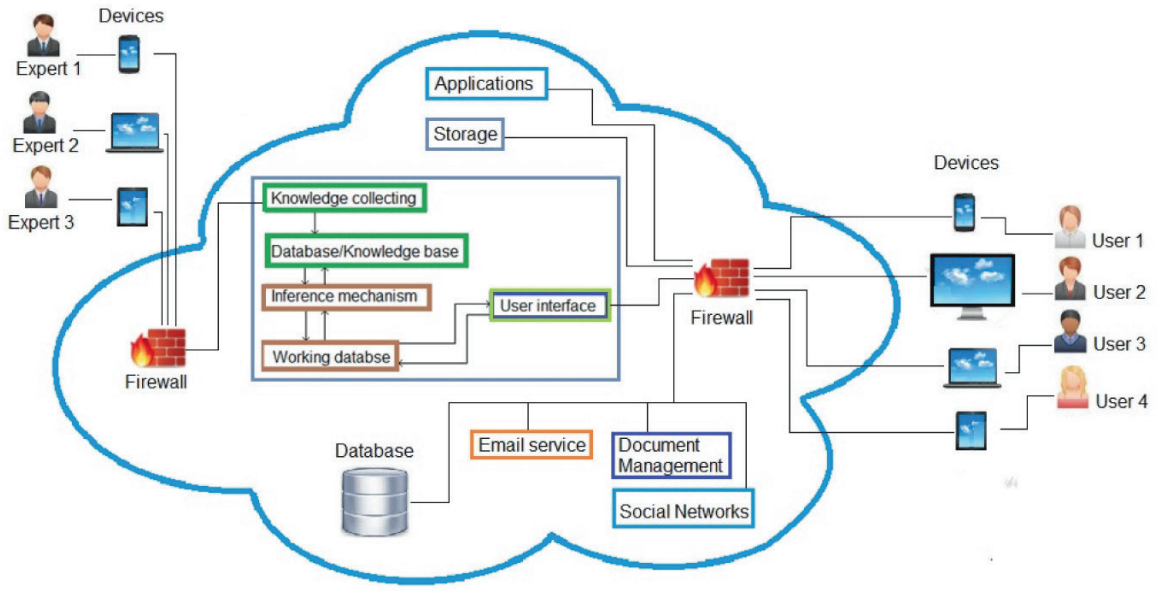



***Figure 3.*** *Cloud-Based Expert System CBES model*
***(Source: own)***

ments offer resources through the internet so basically the main expenditure is the acquisition of devices through which the cloud platform is accessed. Increased collaboration is manifested as free, fast and effective file, data and profile sharing among the users of the cloud platforms.

Furthermore, working from anywhere is enabled through the Internet. Users can access crucial files at the right place and at the right time. Security protocols "defend" the user from external threats. In addition, cloud computing offers easy document control. Users can access, create, share, distribute and publish documents to other users who have authorization to access those documents. Security is indeed important, thus cloud computing environments offer a secured database where all files are stored, and in the situation of a malfunction of the device from which the cloud is accessed, the cloud simply stores the data, and there are no losses.

Competitiveness is presented as the possibility of SMEs to act faster and more effectively than big enterprises. Dynamic markets ask for fast decisions, and cloud computing offers exactly that. A dynamic information distribution gives necessary support for decision making. In addition, cloud computing platforms can serve more clients with the same infrastructure, while the traditional infrastructures are bordered only to the company that implemented them. Therefore, using cloud computing services is more environmental friendly.

Security risks and threats are present in the cloud computing environment. These security risks should be addressed through standardized security and risk management processes [20] [21]. The mentioned research papers [20] [21] are focused on risk management and similarly, cloud computing environments contain a protocol and precise calculations for risks and security threats.

In the next section, the CBES model will be presented. The section includes the scheme of the model, the necessary infrastructure, elements, users and application in practice.

## CBES model

The Cloud-Based Expert System (CBES) model is a platform in a cloud environment where an expert system is implemented with the main purpose for fast-paced decision making. The model includes an expert system, applications, storage options, document and email management, database beside the ones in the expert system and social network connectivity. Furthermore, firewalls are implemented as a form of protection from unauthorized access and security risks and threats. The model is shown in Figure 3.

Figure 3. shows the CBES model. The expert system's knowledge collecting element is connected

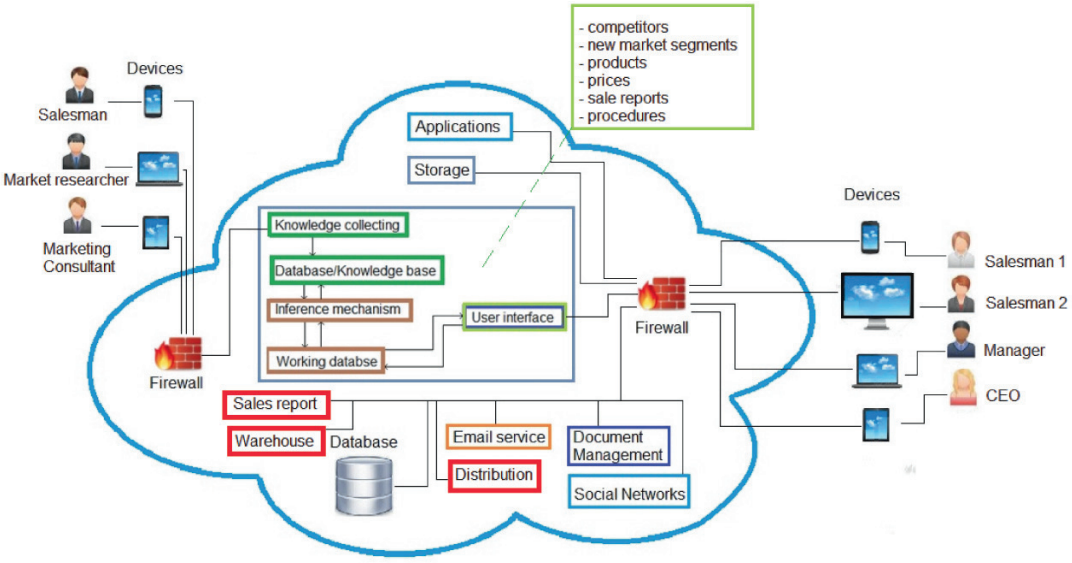



***Figure 4.*** *CBES model on the example of a printer selling company*
***(Source: own)***

with a firewall. Further, an expert or more experts through internet connection and their devices can store knowledge into the knowledge colleting element of the expert system. This, however, does not mean that there is no starting knowledge in the expert system, in the contrary, the expert system should be "ready to go", but additional information and data can be added/stored by the expert through their devices. The firewall secures the connection, therefore not allowing unauthorized persons access to the knowledge collecting element of the expert system. The whole expert system is in a cloud computing environment.

The users access the cloud with their devices through the Internet. Similarly to the experts' access protocol, the users also connect to the cloud computing environment through a firewall. This firewall ensures security from threats and unauthorized log-ins.

Furthermore, the cloud environment contains a database separated from the expert system, an E-mail service, document management applications, social network profiling and additional storage. These resources enable the users to communicate important information and share crucial data for decision making.

The basic concept of the model is that the expert system contains knowledge about structural and fresh information about the market, about the products, prices, sale reports, documents, procedures, time necessary for a procedure completion and the hierarchical structure of the organization or company. When there is a rapid and fast development of new trends, the expert system is storing new info from the experts. In addition, the users can acquire information from other sources and together with the expert's system diagnose a fast, good and effective decision can be made.

The implementation process is not different from other cloud deployment models. The cloud environment can be private, public or a hybrid cloud. An expert system can be implemented in any of the mentioned deployment models. A crude example of the implementation can be shown on a company that sells printers. The CBES scheme for this kind of company is shown in Figure 4.

Figure 4. shows how a cloud-based expert system can be used for fast-paced decision making in a printer selling company. The users are the salesmen, managers and the CEO. The cloud environment includes applications, storage, sale report software, warehouse inventory software, distribution planning, e-mail services, document management, social network integration, firewalls and the expert system.

The experts are or can be salesmen, market researchers or marketing consultants. The information in the knowledge base of the expert system contains information about the competitors, new market segments, products, prices, sale reports and

other procedures. When there is a turbulent atmosphere on the market and the management needs to make a decision about a new shipment to a specific market segment, the expert system can diagnose the best option and outcome of a certain shipment. In addition, the salesmen and manager and CEO as well, can, based on their personal experience and the expert's system diagnosis, make a good decision and define further actions.

Through the cloud computing environment and expert system support, the decision making process is drastically reduced. Costs are reduced and information distribution is more effective compared to traditional information distribution techniques [22] [23] [24].

## Conclusion

The facts presented in this paper clearly show that IT technologies are the future in communication and long distance services. Cloud computing environments will become a crucial part of organizations. Expert systems are a great addition to the business "arsenal" of companies in specific fields. Furthermore, it can be concluded that expert systems in a cloud environment provide a fast and effective support for decision making. In addition, the purpose of an expert system in a cloud environment can be different, and it depends on the function of the cloud. If medical diagnosis is needed, an expert system could be implemented into a cloud platform which is used in a hospital. Reduced costs, flexibility, scalability and security are more than enough reason to consider expert systems and cloud computing environments for business decision making and/or other functions. In addition, cloud platforms offer stable resources and stability.

The CBES model is not an innovative technology, rather than an innovative approach to a developed and already widely used technology. Organizations and enterprises, as well as SMEs, can benefit from this type of cloud deployment. For further research, it is recommended to address companies who had implemented expert systems and/or cloud platforms and analyze the company's improvement first hand. Until then, the CBES model stays as feasible and effective model.

## References

[1] Bahrami, M., & Singhal, M. (2015). *The role of cloud computing architecture in big data. In Information granularity, big data, and computational intelligence*. Springer International Publishing.

[2] Cao, Z., Lin, J., Wan, C., Song, Y., Zhang, Y., & Wang, X. (2016). Optimal Cloud Computing Resource Allocation for Demand Side Management. *IEEE Transactions on Smart Grid,* 99, 1-13

[3] Castillo, E., Gutierrez, J. M., & Hadi, A. S. (2012*). Expert systems and probabilistic network models.* Springer Science & Business Media Rittinghouse, J. W., & Ransome, J. F. (2016). *Cloud computing: implementation, management, and security*. CRC press.

[4] Castillo, E., Gutierrez, J. M., & Hadi, A. S. (2012). *Expert systems and probabilistic network models*. Springer Science & Business Media.

[5] Chang, V., Kuo, Y. H., & Ramachandran, M. (2016). Cloud computing adoption framework: A security framework for business clouds. *Future Generation Computer Systems,* 57, 24-41.

[6] Chen, Y., Hsu, C. Y., Liu, L., & Yang, S. (2012). Constructing a nutrition diagnosis expert system. *Expert Systems with Applications*, 39(2), 2132-2156.

[7] Chen, Y., Hsu, C. Y., Liu, L., & Yang, S. (2012). Constructing a nutrition diagnosis expert system. *Expert Systems with Applications*, 39(2), 2132-2156.

[8] Chu, J., Chu, J., Wu, J., Wu, J., Zhu, Q., Zhu, Q., ... & Sun, J. (2016). Resource scheduling in a private cloud environment: an efficiency priority perspective. *Kybernetes,* 45(10), 1524-1541.

[9] Guo, F., Yu, L., Tian, S., & Yu, J. (2015). A workflow task scheduling algorithm based on the resources' fuzzy clustering in cloud computing environment*. International Journal of Communication Systems,* 28(6), 1053-1067.

[10] Kaufman, L. M. (2009). Data security in the world of cloud computing. *IEEE Security & Privacy*, 7(4).

[11] Kidd, A. (Ed.). (2012*). Knowledge acquisition for expert systems: A practical handbook*. Springer Science & Business Media.

[12] Kiritsis, D. (1995). A review of knowledge-based expert systems for process planning. Methods and problems. *The International Journal of Advanced Manufacturing Technology,* 10(4), 240-262.

[13] Lisenkov, V. M., & Bestemyanov, P. F. (2015). Safety and Risk Management. *Journal of Information Technology & Applications*, 5(1), 42-55.

[14] Marcus, S. (Ed.). (2013). *Automating knowledge acquisition for expert systems (Vol. 57)*. Springer Science & Business Media.

[15] Petrović, I., Bočvarski, V., & Petrović, V. (2016). Expert systems in physics: Methodology and application. *Physics Essays*, 29(1), 49-56.

[16] Rittinghouse, J. W., & Ransome, J. F. (2016). *Cloud computing: implementation, management, and security*. CRC press.

[17] Rybina, G. V., & Blokhin, Y. M. (2016, September). Use of intelligent planning for integrated expert systems development. In Intelligent Systems (IS), *2016 IEEE 8th International Conference on*, 295-300.

[18] Sellami, R., Bhiri, S., & Defude, B. (2016). Supporting multi data stores applications in cloud environments. *IEEE Transactions on Services Computing*, 9(1), 59-71.

[19] Shortliffe, E. (Ed.). (2012*). Computer-based medical consultations:* MYCIN (Vol. 2). Elsevier.

[20] Tograph, B. & Morgens, Y. R. (2008). Cloud computing. *Communications of the ACM*, 51(7), 9-11.

[21] Wagner, W. P. (2017). Trends in expert system development: A longitudinal content analysis of over thirty years of expert system case studies. *Expert Systems with Applications,* 76, 85-96.

[22] Zimmermann, H. J. (2012). *Fuzzy sets, decision making, and expert systems (Vol. 10).* Springer Science & Business Media.

[23] Zhong, Z., Chen, K., Zhai, X., & Zhou, S. (2016). Virtual machine-based task scheduling algorithm in a cloud computing environment. *Tsinghua Science and Technology*, 21(6), 660-667.

Submitted: March 20, 2017.
Accepted: May 28, 2017.

## About the Authors

**Mihalj Bakator,** MSc Mihalj Bakator was born in Pančevo, Serbia in 1992. He received the MSc degree in Engineering Management at the University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin in 2015. Currently, he is a PhD student of engineering management in Zrenjanin. His areas of interest are: cloud computing technologies, marketing

**Dr Dragica Radosav - full professor,** graduated in June 1980. She defended her M.A. thesis in 1991. She defended her doctoral dissertation (title: Stuff education research for systems with integrated document management) in 1994. She has had a permanent employment at the Technical Faculty "Mihajlo Pupin" since 1983 as an assistant, and from 1994 as a teacher she has been engaged in a great number of subjects. During the past years on undergraduate study level, she has been teaching the subjects as follows Information Technology, Software Engineering1 and Software Engineering 2. On graduate study level, she has been teaching Information System Design, Geographic Information Systems and Methodology of Pedagogical Research and Methodology of Scientific Research Work. She teaches the subject Methodology of Scientific Work on doctoral study level. Ms Dragica Radosav is author/co-author of 11 textbooks. During the last five years she published 9 papers in Journals from the SCI list (M23), i paper in the cathegory M21, 5 papaers in the cathegory M24, 79 papers on international conferences and 12 papers in international and scientific journals and 38 papaers on national conferences. She has puplished total 12 textbooks that are used in teaching of many subjects. She was reviewer of 14 textbooks.

# MODELLING THE PROCESS OF IS AUDITING IN THE PUBLIC ADMINISTRATION USING UML DIAGRAMS

**Dalibor Drljača[1], Branko Latinović[2], Dušan Starčević[2]**

*[1]Europrojekt centar, drljacad@gmail.com, [2]Pan European University Apeiron Banja Luka*

*branko.b.latinovic@apeiron-edu.eu, dusan.b.starcevic@apeiron-edu.eu*

**Case study**

DOI: **10.7251/JIT1701032D** UDC: **35.07/.08:004.438UML**

**Abstract:** Although information system audit is a very important business process, at present this is not obligatory in the public administration institutions in Bosnia and Herzegovina and Republic of Srpska. Due to the importance of this process, this paper proposes a model for auditing of information systems in the public administration institutions. The model intends to explain the audit process using a visual representation of the process with UML diagrams. UML is an internationally recognised language for business process modelling and has a number of advantages over other similar languages and standards. Therefore, UML is selected in modelling for modelling of information system auditing process in the public administration institutions.

**Keywords:** UML, auditing, information systems, public administration, business process modelling.

## INTRODUCTION

Implementation of e-government strongly depends on implemented and engaged information system with corresponding applications. In order to provide e-services to citizens and business, these information systems are dealing with a lot of sensitive data, such as bank accounts, transactions of funds, private data, etc. Due to the importance of these data, the public administration institutions need to secure communication and transaction of such data. For this, they use a number of different methods and tools. The information systems are very specific and require more attention in the provision of safety and security. The audit of information systems recently became a very important business process in companies and institutions dealing with sensitive data, such as investment funds, corporations, banks, etc. However, in Bosnia and Herzegovina, there is no legal obligation for the introduction of information system audit in the public administration institutions at any level (municipal, cantonal, entity or state level).

This paper intends to offer a model of information system audit in the public administration institutions using Unified Modeling Language (UML) diagrams. There are different types of public administration institutions and this model takes as an example an auditing process in one Ministry. The differences in the model may occur depending on the type of the institution, but these differences are minor. For example, in the Ministry, the ultimate authority is Minister, while in some Agencies this authority is the Director, etc. But this model does not depend on the size of the institution, which is more important. Also, the model assumes that the audit is performed by the external auditor, due to the fact that this is not legally regulated.

## METHODS AND MATERIALS

The paper deals with several scientific fields, but the predominant are economics (in the frame of fi-

nances and auditing), information technologies (in the frame of information systems, structure and security), and law (in the frame of legislation and public administration structure and operations). Therefore, the methods and methodologies selected have to respond to such scientific mix – they **combine modern methods and tools (typical for information technologies) and traditional ones (typical for economics and law)**. This mix of methods and tools **enables a holistic approach to solving the problem**. Also, the basic research methods (applicable for all scientific fields) such as **deduction and induction, synthesis and analysis**, are used for better structuring of conclusions.

The main problem was **scarce literature available in local language** for the topic of auditing of the information system. The problem of information system auditing is not adequately presented in scientific or educational literature in Bosnia and Herzegovina and it is possible to conclude that such literature does not exist at all. There are just some individual attempts and efforts invested in the publishing of articles dealing with the auditing of information systems. Such articles are even more narrowed to an area of cyber-security and how to protect information systems, but not to auditing as the process, nor the use of UML in the modelling of such processes. Therefore, **the use of Internet as the primary source for most adequate information** was selected to compensate this lack of literature. As the auditing of information systems is very specific and information-sensitive process, the most important providers of standards for auditing were consulted such as ISACA (Information Systems Audit and Control Association), ISO (International Organization for Standardization) and INTOSAI (International Organisation of Supreme Audit Institutions). **Use of international standards** is important for maintaining the objectivity of the auditing process.

Especially, it is important to use **normative methods** to discover and define norms and rules of the process. As the paper deals with information systems in public administration, it was necessary to also use the **cybernetic method** to explore the functionality of complex public administration systems and its management. Most of the literature for investigation of this was found in the legal acts of the public administration institutions, Official Gazette and other legal documents. However, in order to make a generally accepted model, the **generalization method** was used to generalize and systematize acquired knowledge. **Modelling method** is frequently in use in the software engineering. This method allows consideration of all features, aspects and the impact of the business process being modelled.

### Languages for modelling of business processes

In order to make a model of the IS auditing, for modelling of auditing as the business process we needed adequate modelling language. Each of modelling languages has its own syntax (set of graphics for description – notations) and semantics (interpretation of these graphic notations – the meaning) [1] The modern science recognizes a number of these languages, but there are three most accepted:

- **Business Process Modelling Notation** (BPMN),
- **Unified Modelling Language** (UML), and
- **Event-driven Process Chain** (EPC) or **extended Event-driven Process Chains** (eEPC).

**BPMN** is the international standard for modelling of business process maintained by the *Object Management Group* (OMG). The aim of BPMN is to provide a notation that integrates the best business process modelling practices and enable a better understanding of the notation by business analytics, programmers and business users [2]. BPMN uses simple diagrams that are created with the **limited set of the graphic elements** grouped into:

- **Flow objects** (*Events, activities, gateways*);
- **Connecting objects** (*Sequence flow, message flow, association*);
- **Swimlanes** (*Pool, lane*); and
- **Artefacts** (*Data object, group, annotation*).

BPMN business model is realized through **three sub-models**:

- **Processes or Orchestration** including *Private non-executable (internal) Business Processes*, *Private executable (internal) Business Processes*, and *Public Processes*;
- **Choreographies**, and
- **Collaborations**, which may include Processes or Choreographies.

**UML** is language and standard maintained also by OMG. But, it is universal and has more extending capabilities than BPMN. UML is platform-independent, universal modelling language primarily made for the creation of software solutions. [3]. Due to its simplicity and universality, it became a part of modern CASE (*Computer-Aided Software Engineering*) tools. Since its standardization in 1997, it significantly improved software development but also information systems development. In the core of UML lies object-oriented paradigm (OO) for software development, originating from three different OO development methods:

- *Object-Oriented Design OOD,*
- *Object-Oriented Software Engineering OOSE* and
- *Object-Modeling Technique OMT*.

According to the latest standard version (2.5 from March 2015), UML notation contains set of graphical symbols, but significantly larger set than the one in BPMN, and grouped into:

- ***Things*** (for description of behaviour, structure, grouping and explanation),
- ***Relations*** (main four: dependence, association, generalization and realization),
- ***Diagrams*** (7 structural and 8 behavioural)
- ***Mechanisms*** (general mechanisms and extensibility mechanisms)

**EPC** was created in 1992 in Germany by the Institute for Information Systems in Saarbrücken and World-famous company SAP [4]. Although developed before some other languages and very intensive in use, even today EPC is not standardized [5]. The notation uses following elements:

- **Event**,
- **Function**,
- **Process owner**,
- **Organisation unit**,
- **Information**, **material**, **or resource object**,
- **Logical connector,**
- **Logical relationships** that include,
  - Branch/Merge – XOR operand,
  - Fork/Join – AND operand,
  - OR operand,
- **Control flow**,
- **Information flow**,
- **Organisation unit assignment,**
- **Process path**.

Soon upon establishment of EPC, many shortcomings were observed, which caused the creation of **eEPC** (**extended EPC**), but also a dozen of other extensions of the basic and original EPC that prevented its standardization. **eEPC** is the main business process modelling language in ARIS (Architecture of Integrated Information Systems) managed by IDS Scheer AG from Germany.

Although all of three selected languages have their advantages and shortcomings, it is up to the business analyst to choose the adequate one for the task of system analysis. Having in mind possibility for automation of the whole process, UML was selected as the language that supports many CASE tools and with the ability to convert diagrams into the programming code. Also, one of the main reasons for selecting UML is in its standardized approach, universality, and the ability to document processes important in the last phase of SDLC (Software Development Life Cycle). For the presentation purposes of this paper, Use case diagrams, Activity diagrams and Sequence diagrams will be mostly used. Also, we will consider the whole auditing process as a system for itself and therefore, we will use Use Case diagrams to present general scenarios of the system›s functionalities, and Activity and Sequence diagrams to described flow of the processes within the system. All these diagrams aim to present „what" system should do focusing on the role of the user and not on the implementation and realization details [6]. As CASE tool, which supports UML, the preference was to use a Visual Paradigm Community Edition v.14 (https://www.visual-paradigm.com). This Community Edition has all necessary functionalities and it is a freeware.

## Results and discussion

According to existing legislation, the public administration institutions in Republic of Srpska and Bosnia and Herzegovina are not obliged to make a regular audit of the information systems in use. However, there is a lot of information systems in the public administration that were acquired without considerations about interoperability of these systems.

Therefore, for the purpose of modelling such process, a hypothetical approach is made using present organisational and functional structure of the public administration institutions that is given in the relevant legislation. The model considers a case of the information system in the Ministry, as it is the institution that is dealing with sensitive and important data requiring safe and functional information system. Also, the model considers the case of the auditing process done by the external auditing company since there are no regulations dealing with this issue.

### Description of the IS audit process in the public administration institutions

The whole process has to start from the top-management and in the case of the Ministry, it is the Minister who should initiate this process. This process can be initiated based on legal obligations or voluntarily if the Minister considers that the security and safety of the data and information system are endangered. Again, there is no legal requirement and the model considers voluntary initiation.

In order to describe the process, a number of actors are recognised in the system that is of crucial importance for the realization of the process.Without these actors, the auditing (as the system) is not operable:

- The Minister (Minister)
- The legal department personnel (LegalDept)
- The financial department personnel (FinancDept)
- The employees of the Ministry (Empl)
- The external auditor personnel (ExtAud)

The actors are generalized for description purposes and there can be several executors within one actor (except for the Minister). Also, other actors can appear in the process and their role will be explained from case to case.

In brief, the description of the actors' roles is as follows:

- **The Minister** (as supreme authority) initiates the auditing process and initiates the procurement of the external auditor that will perform auditing process.
- **The legal department** in the Ministry is in charge of preparing all necessary documentation – both for procurement of the external auditor and for goals of the audit (Audit Charter) that will be at the same time the work description and content of the contract with the external auditor.
- **The financial department** task is to service all procedures related to the financial arrangements of the process (payments)
- **The employees** in the Ministry are expected to take part in the auditing process in accordance with the planned use of resources for the audit. It may be the case where the employees will play an active role (answering the questions, providing some tasks, etc.) or the case where they will be in a passive role (they will be looked as a part of the whole team or the system).
- **The external auditor** is expected to respond to the procurement call and it will include preparation of the procurement materials and applying process. Also, upon signing of the contract, external auditor should prepare the auditing conditions and goals together with the representatives of the Ministry (Legal dept). After that, auditor will perform the core of the audit and at the end will produce and deliver the auditing report that has to be agreed with the Ministry

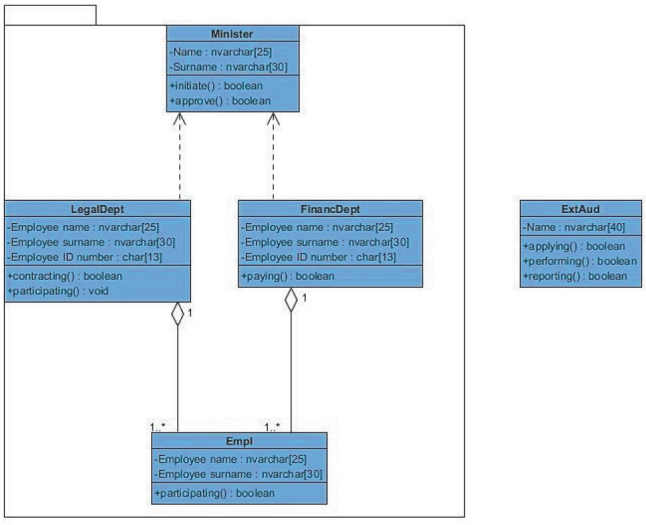



***Figure 1.*** *Class diagram for IS audit system (author)*

Four main business processes are identified (functionalities) that are presented with UML Use

case diagrams:

- Audit initiation,
- Procurement of the external auditor,
- Auditing process,
- Reporting.

Therefore, the context diagram presented with UML Use case diagram looks as in illustration 2 below.

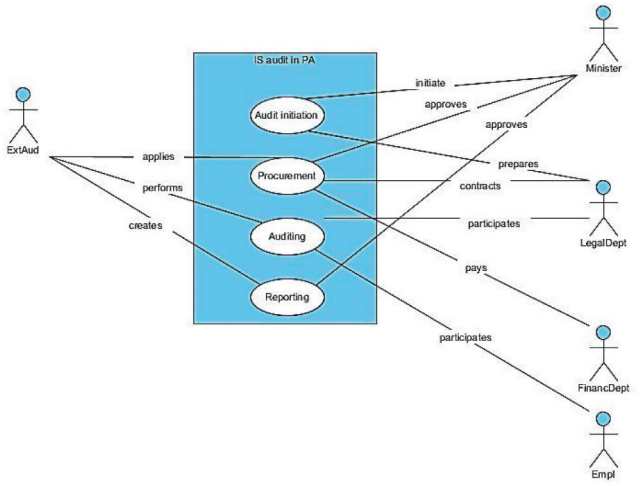



***Figure 2.*** *Context diagram presented with UML Use Case diagram (author)*

The whole system is divided into four main functionalities at the context level and it is presented with the UML Use Case diagram (Figure 2.) titled „IS audit in PA". This diagram shows main interactions between identified actors that are responsible for planning, organizing, performing and reporting in the auditing process. The ExtAud actor is on the left side (usually the side of system's user), while other actors are on the right-hand side of the system (usually the side of system's owner and maintenance). These functionalities should lead to the completion of the auditing process and are in line with the local legislation regulating work of the Ministry.

The first functionality - „**Audit initiation**" covers set of activities and procedures needed to initiate auditing process. Unless it is regulated in a different way, the initiator should be institution authorized person – in this case, actor „Minister". We presumed this case, but it can be the case (such as in banking system) that the auditing process is initiated automatically and similar to the financial audit. It the latter case, there is a slight and not essential change of the model. It is very crucial that this functionality is initiated by the top-level authority since information system audit will have at disposal all elements of the information system (hardware, software, employees, network, etc.) in order to perform check of control points. This functionality engages actors „Minister" and „LegalDept" as the leading actors.

The flow of activities is following (as presented in Figure 3):

- „Minister" initiates (orders) the audit process by sending formal request to the „LegalDept";
- „LegalDept" examines the internal procedures, collects materials and prepares necessary documents for „Minister"'s decision;
- After preparation of the legal document, „LegalDept" sends the documents to the „Minister" for signature;
- „Minister" signs the official documents and returns it to the „LegalDept".

Having the documents signed, the „LegalDept" can now close the functionality „Audit initiation" and start the functionality „Procurement" of the external auditor.

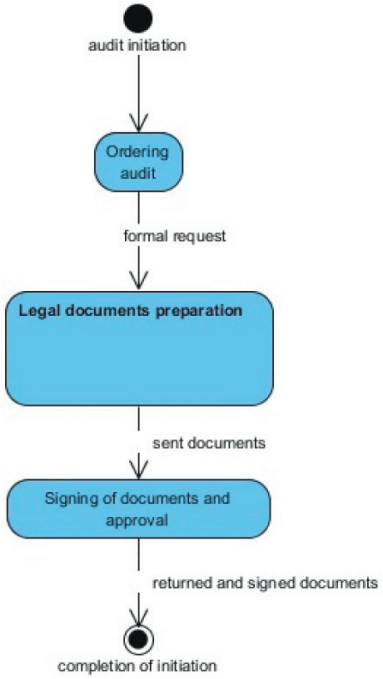



***Figure 3.*** *Activity diagram for audit initiation (author)*

The functionality „**Procurement**" aims to make procurement of the external auditor for information system auditing. The prerequisite for this functionality is the completion of the functionality „Audit initiation" as described above. This functionality is

not necessary if the auditing process is performed by the internal auditor. Since there is no explicit regulation (actually no regulation at all), we assumed procurement of the external auditor. As a part of the previous functionality, the „LegalDept" should create one document with details on the content of the audit. This will use as the description of the work and will constitute the Audit Charter. It is a crucial document because it expresses the wish of the institution to do the audit and what volume of the audit, which controls to check and test, what resources will be analyzed etc. As this document is a „wish" of the institution, the agreement can be made upon selection of the auditor to extend or modify the content and framework of the audit. However, this should be done in a mutual agreement since the content and framework shall determine the price of the service and this should be done prior to the contracting. This functionality is split into two parts:

- First part is commercial procurement and procedures related to the procurement (tender). This part will end but the functionality then goes to the stand-by, and
- The second part of this functionality will start after the performed audit and relates to the payment procedure for completed task.

The flow of activities is following (as presented in Figure 4):

- „LegalDept" based on official documents prepares the call text for the advertisement and initiates procurement procedure by sending the add to the „Media"
- „Media" is publishing the add, and sending the invoice to the „LegalDept"
- „LegalDept" forwards the invoice to the „FinancDept" that prepares documents for payment.
- „FinancDept" forwards payment documents to the „Minister" for signature and approval.
- „Minister" signs the documents and returns it to the „FinancDept" that makes the payment to the „Media".
- During the advertising period (declared by the legislation), the auditing companies submit application („ExtAud")
- „LegalDept" makes a proposal for the selection „Committee" and forwards it to the „Minister" for approval.

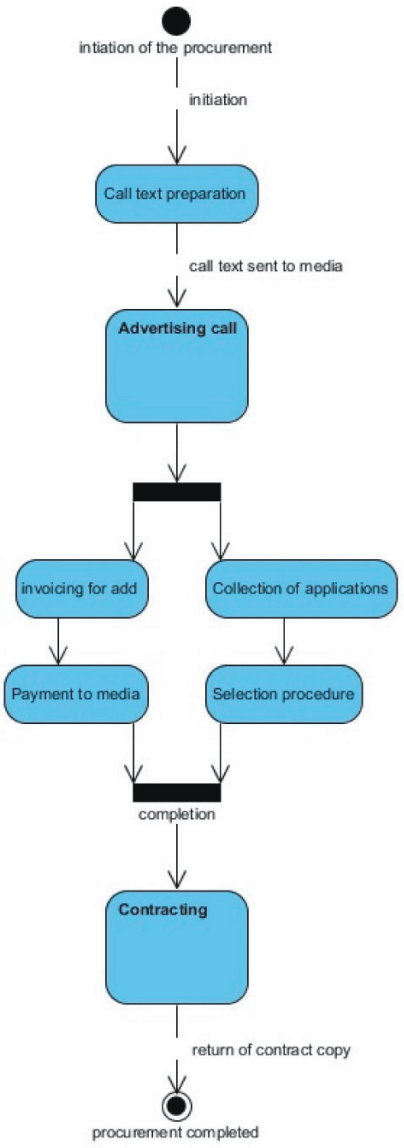



***Figure 4.*** *Activity diagram for external auditor procurement (author)*

- „Minister" approves and appoints the „Committee" that will review all applications and make a selection of the most suitable bidder.
- „LegalDept" is in charge of the whole selection process and prepares the final decision on selection based on the report from the committee and forwards it to the „Minister" for approval.
- „Minister" approves the selected company and orders contracting procedure from the „LegalDept"
- „LegalDept" creates a contract and forwards it to the selected company for signature.
- If all clauses are ok, „ExtAud" signs the contract and returns it to the „LegalDept"

- „LegalDept“ forwards the contract to the „Minister“ for signature.
- „Minister“ signs the contract and returns it to the „LegalDept.
- „LegalDept“ dispatches the copy of the contract to the „ExtAud“ and with this, the procurement process is completed.

Upon receipt of the copy of the contract, „ExtAud“ is starting its own preparations for the audit (selection of staff, preparation of template materials, etc). „ExtAud“ is now ready to schedule the whole auditing knowing what exactly it has to do. Therefore, the first step in new functionality called „**Auditing**“ is organizing an initial meeting to fix the dates for a physical visit to the site and to the information system. During this meeting, the last details are negotiated for smooth and seamless audit. „ExtAud“ introduces working templates, explains the flow of action and fixes other dates for visit, if necessary. It is important that in this meeting all key stakeholders in the process participate, especially the „Minister“ as it has to authorize the use of resources (hardware, software, access, employee etc). After the meeting, „ExtAud“ is fully ready to start the auditing process. This process has its dynamics defined at the meeting. If everything is done smoothly, the auditing process ends with „ExtAud“ last meeting organized with „Empl“ (employees on site). For performing the audit, the „ExtAud“ shall use methodology prescribed by some of internationally accepted standards and frameworks for an audit of information systems, such as COBIT, ITIL, ISO, VAL-IT etc.

The flow of activities is following (as presented in Figure 5):

- „ExtAud“ receives the copy of signed contract and starts its internal preparations;
- „ExtAud“ proposes the term for the first joint meeting for organization of work;
- The meeting is organized and realized, all details are agreed;
- „ExtAud“ performs activities within auditing process with assistance from „Empl“;
- If needed, „ExtAud“ calls for another joint meeting to finalize the work;

Otherwise, with its last visit „ExtAud“ completes the auditing process.

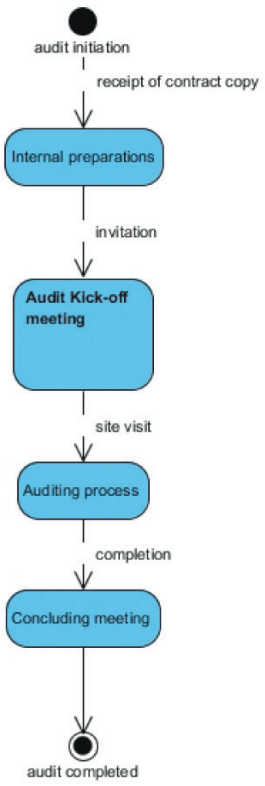



***Figure 5.*** *Activity diagram for auditing process (author)*

Now the auditing process is complete and functionality „**Reporting**“ starts. „ExtAud“ is back to its premises and organizing the work for preparation of the audit report. The audit report has own structure and has to be completed in agreement with the institution ordering the audit. This means that „ExtAud“ has to prepare the first draft of the report and discuss it with the institution. This will be done at one joint meeting organized in the premises of the institution and it is very similar like in the case of a financial audit. „ExtAud“ presents the findings and discusses them with the management of the institution and key stakeholders that participated in the auditing. This is positive practice because it may help „ExtAud“ to clear some suspicious elements which occurred during the preparation of the report. Upon the completion of this meeting, „ExtAud“ is back home and formulating the final version of the report, including the invoice for the services. At this point, we are back in the functionality of „Procurement“ – second part dealing with payment of auditor’s work.

The flow of activities is following (as presented in Figure 6):

- „ExtAud“ is working on a report in its own premises;
- „ExtAud“ drafts report and organizes the meeting in the institution;
- The meeting is done in the institution and draft report is analyzed with key actors;
- „ExtAud“ completes the report in its own premises and sends the final report with the invoice;
- „FinancDept“ completes the payment.

Upon payment completion, the whole „Reporting“ and „Procurement“ functionalities are completed, as well as the whole proposed auditing process.

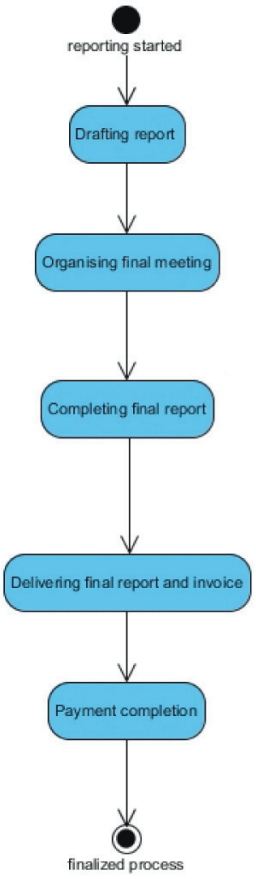



***Figure 6.*** *Activity diagram for reporting functionality (author)*

Each of this Use Case can be further decomposed due to the complexity of the procedures, and one proposal is given in figure 7 in a form of UML Sequence diagram. The Sequence diagram gives an overview of the process as a whole divided into frames in accordance with the described four functionalities.

## Conclusion

The auditing of the information systems is becoming a crucial business process for institutions and companies dealing with a large quantity of very sensitive data, such as banks, insurance companies, telecoms etc. However, information systems in the public administration institutions should also be a subject of the regular audit. Due to the importance of data in such information systems, they can be a target of misuse that should be prevented from the initiation of the information system.

The intention of this paper was to present one of the possible models for auditing of the information system in public administration institutions. The proposed model can have slight changes and variations due to the complexity and internal organization of the institution. This model presented auditing in the Ministry, where the ultimate decision-making authority is the Minister. The model follows four (4) main functionalities that can be decomposed further on individual actions, even to the level of automation of the process. There were two options for engagement of the auditor: internally or externally. Again, the model could be slightly changed, but it was decided to present the case of engagement of the external auditor. There are a number of reasons why the external auditor should be preferred. One of them is that the external auditor can have more experience, knowledge and expertise in implementing the auditing of the information system compare to the internal auditor. The public administration institutions usually do not have such profile among the employees. Also, to acquire the psychological effects of the audit and to obtain objective and relevant auditing data, the engagement of the external auditor justifies this decision for described model.

For the visualisation and description of the modelling process UML was used as a worldwide recognised unified modelling language. The main classes or actors were presented using the Class diagram. A context diagram of the auditing was given using Use Case diagram showing the top-level functionalities of the auditing system. More details on the flow of the auditing process were given with the use of Activity diagrams for each phase of the process. Overall functioning of the system and relevant steps were presented with a Sequence diagram presenting the system functionalities against the flow of the time.

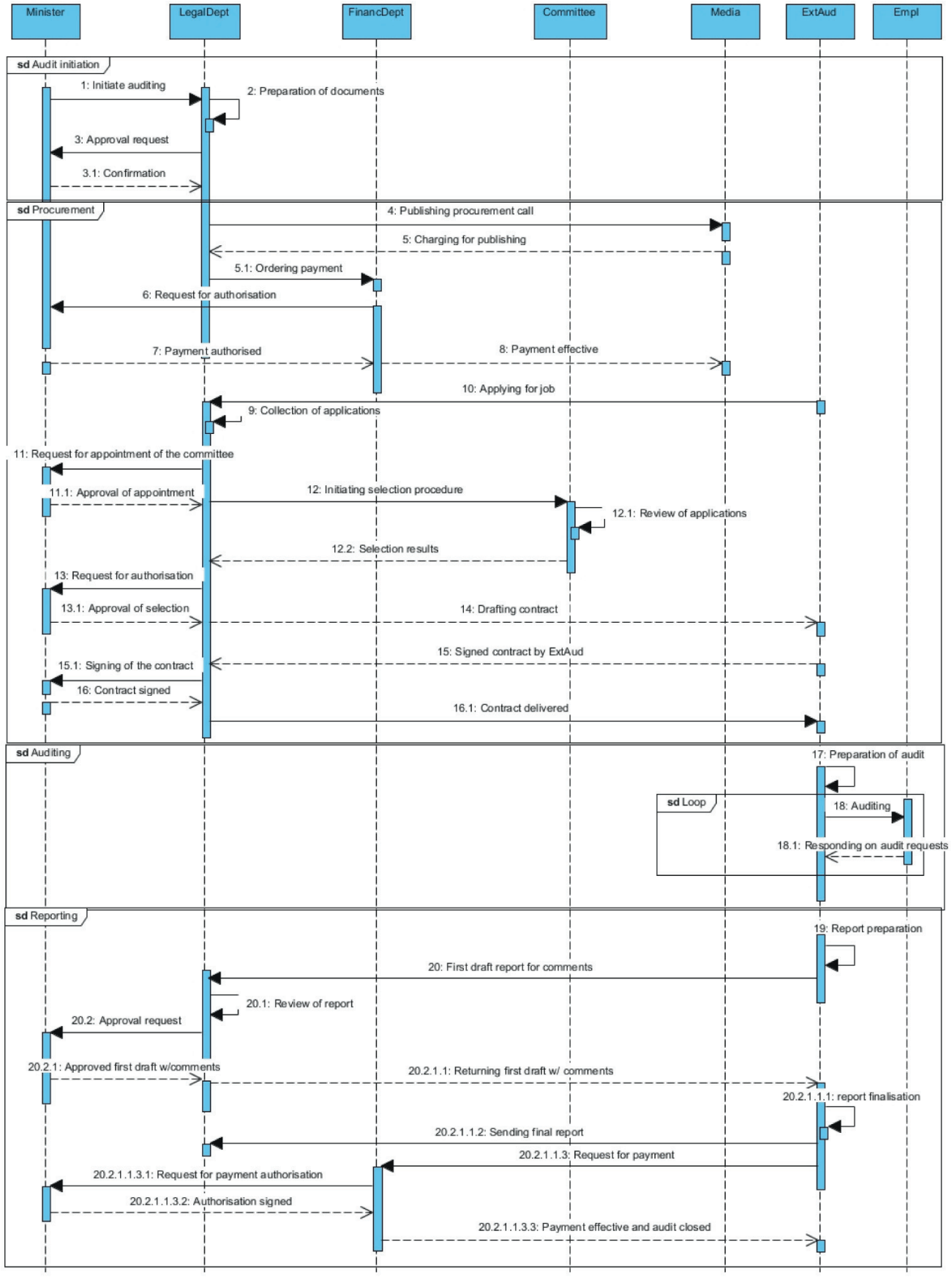



***Figure 7.*** *UML Sequence diagram presenting flow of procedures in auditing process (author)*

With its essence, this paper is a pioneering step in Bosnia and Herzegovina and even abroad. The novelty is in the first instance, a proposing of the model since the auditing of the information systems is (in the contemporary literature) described just textually with the use of recommended and recognised frameworks and standards. It was always up to the auditing body (usually a non-IT expert) to define this process. This leads to the second novelty of this model and that is the use of CASE tools (in this case Visual Paradigm) for modelling and visualisation purposes.

This way the auditing process is seen from the perspective of IT and opens a serial of new questions related to the improvement of the process, proposed model but also wide-recognized standards for the auditing of the information systems. In the future, it is necessary to further develop proposed model to the level of automatization and possible computerization of the process. The auditors are using CAATT (Computer Aided Techniques and Tools) of a different kind for data acquisition and processing, in accordance with the its own preference of the auditor. One of the possible new insights could be the influence of selected CAATT tool for success and quality of the auditing process.

## References

[1] Harel D, Rumpe B (2004) Modeling Languages: Syntax, Semantics and all that Stuff, The Weizmann Institute of Science, Israel

[2] Object Management Group (2011) Business Process Model and Notation (BPMN) Version 2.0, OMG Document Number: formal/2011-01-03, Standard document URL: http://www.omg.org/spec/BPMN/2.0

[3] Object Management Group (2015) OMG Unified Modelling Language (OMG UML) Version 2.5, OMG Document Number: formal/2015-03-01, Standard document URL: http://www.omg.org/spec/UML/2.5

[4] Keller G, Nüttgens M, Scheer AW (1992) Semantische Prozeßmodellierung auf der Grundlage - Ereignisgesteuerter Prozeßketten (EPK) In: Veröffentlichungen des Instituts für Wirtschaftsinformatik (IWi), No. 89, Universität des Saarlandes

[5] Ko RKL, Lee SSG, Lee EW (2009) Business process management (BPM) standards: a survey, Business Process Management Journal 15, pp. 744–781.

[6] Pilone D, Pitman N (2005) UML2.0 in a Nutshell, O'Reilly Media, Inc., USA

Submitted: April 7, 2017.
Accepted: June 13, 2017.

## About the Authors

**Dalibor Drljača** is a Ph.D. candidate at the Faculty of Information Technologies at the Pan-European University APEIRON Banja Luka and has MA in information technology and MA in technologies for the Development of European Projects. His main research interests are in e-Government, audit of information systems and e-Commerce. He is part-time engaged as a Senior teaching and research assistant at Pan-European University APEIRON Banja Luka.

**Branko Latinović**, Ph.D., is a full-time professor and Dean of the Faculty of Information Technologies at the Pan-European University APEIRON Banja Luka since its establishment. His research interests are in information systems, e-Commerce and e-Government.

**Dušan Starčević** is a full professor at the Faculty of Organizational Sciences, University of Belgrade. His main research interests include human–computer interaction, distributed information systems, multimedia, and computer graphics. Starcevic received his BS and MS in electrical engineering and his PhD in information systems from the University of Belgrade.

# COMPARISON OF PERCEIVED INTERACTIVITY MEASURES OF ACTUAL WEBSITES INTERACTIVITY

**Velimir Štavljanin[1], Milica Jevremović[2]**

*[1]University of Belgrade, Faculty of organizational sciences, Belgrade, velimirs@fon.bg.ac.rs*

*[2]School of Electrical Engineering and IT of Applied Studies (VISER), Belgrade, milicag@viser.edu.rs*

**Case study**

DOI: **10.7251/JIT1701042S** UDC: **004.6:004.738.1**

**Abstract:** Interactivity is a concept of enormous importance for digital marketing. It was recognized as a key feature of website, a hub of all digital marketing activities. But, almost all interactivity measures were conceptualized one or two decades ago. In the meantime, technological novelties changed the face of websites. Also, a number of interactivity features increased exponentially. Those changes had a huge impact on practice and could influence user's perception of interactivity. Aim of this paper is to explore whether several selected existing measures of perceived interactivity could cope with those changes. Paper reports a study in which two websites of low and high interactivity were developed and in an experimental setting as stimuli used to test three perceived interactivity measures. Results show that all measures estimated perceived interactivity of a high interactivity website better than of a low interactivity website. Also, results show that particular dimensions of a model could be used to estimate overall interactivity.

**Keywords:** website interactivity, perceptual interactivity, actual interactivity, interactivity measures, website design.

## INTRODUCTION

Interactivity was recognized as the most significant feature of the internet and new media [20, 27]. From the early days of internet, various digital systems have been flooding the market. Interactivity and promises that interactive digital systems offered to users were most important factors that had influenced consumer decision to buy. During 1980s researchers began to study interactivity. Many of them successfully defined interactivity [6, 7, 9, 14, 18, 19, 21, 25, 30, 36], but from different backgrounds and perspectives. Researchers have agreed on the subject that interactivity is complex and multidimensional concept. But there is no consensus about dimensions and elements of interactivity.

Interactivity has numerous positive effects, like acceptance and satisfaction [21]. McMillian [19] stated that interactivity affects the attitude towards the website, the relevance of the topics on the site, returning to the Web site, inviting others on the web site, and purchasing from the website. Interactivity also affects better processing of information on the website and better processing of product information. In order to achieve positive effects in practice, it is necessary to be careful with the implementation. All dimensions of selected model should receive adequate attention. But for practitioners appropriateness of a particular model is always a question. On e more difficulty for practice is that only several measures of interactivity exist.

Paper is organized as follows. First we review selected definitions of interactivity and measures of interactivity. Then we describe details of our research: population, procedure, and stimuli. Next, we present results of our research. Finally, we discuss our research as well as the implication for marketing researchers and practitioners.

## Interactivity

As we mentioned earlier, interactivity has been recognized as the most important feature of the new digital media. However, interactivity soon became synonymous with the Web, so the terms like Web marketing and Web advertising has become an Interactive marketing and Interactive advertising [18]. Mohammed et al. [20] argue that interactivity provides such level of a dialogue that has not been previously known in the history of business. The importance of the concept created enormous interest among researchers.

Rogers [25] provided one of the first definitions of interactivity. He defined interactivity as "the capability of new communication systems (usually containing a computer as one component) to 'talk back' to the user, almost like an individual participating in a conversation".

Rafaeli based his definition on the concept of possible response rate as a measure of media capability to accept and react to responses given to the user, i.e. measure to what extent one message in the exchange is based on previous messages. Rafaeli defined interactivity [21] as "an expression of the extent that, in a given series of communication changes, any third (or later) transmission (or message) is related to the degree to which previous exchanges referred to even earlier transmissions".

Steuer defined interactivity in the context of virtual reality as a determinant of telepresence. Interactivity [30] "is the extent to which users can participate in modifying the form and content of the mediated environment in real time".

Jensen [11] criticized previous practice of linking interactivity with technology. Jensen defined interactivity as "a measure of a media's potential ability to let the user exert an influence on the content and/or form of the mediated communication".

After comprehensive analysis of the technology and communication oriented definitions, Kiousis [14] defined interactivity as "as the degree to which a communication technology can create a mediated environment in which participants can communicate (one-to-one, one-to-many, and many-to-many), both synchronously and asynchronously, and participate in reciprocal message exchanges (third-order dependency). With regard to human users, it additionally refers to their ability to perceive the experience as a simulation of interpersonal communication and increase their awareness of telepresence."

Yadav and Varadarajan [35] defined domain specific definition in the field of electronic marketplaces. Interactivity "in the electronic marketplace is the degree to which computer mediated communication is perceived by each of the communicating entities to be (a) bidirectional, (b) timely, (c) mutually controllable, and (d) responsive."

Johnson, Bruner, and Kumar [13] defined interactivity as "the extent to which an actor involved in a communication episode perceives the communication to be reciprocal, responsive, speedy, and characterized by the use of nonverbal information."

By analyzing all this definitions it is obvious that interactivity is complex and multidimensional concept. Certain definitions describe it as one-dimensional, some as two, three, and four, even as a concept with six dimensions [11]. For practitioners a larger number implies more efforts, and it is easier to analyze and develop particular digital system using less interactivity dimensions. One more difficulty for practice is a fact that interactivity can be actual and perceptual [17, 29, 32, 33, 34]. Actual interactivity, sometimes called structural or objective or feature based interactivity is potential in a medium for interaction [33]. Actual interactivity was operationalized as presence or absence of some interactivity features [29]. Perceived or subjective interactivity can be defined as a psychological state experienced by the user in the process of interaction [33].

## Measuring interactivity

Although many studies have been dealing with modeling of interactivity, only several instruments for measuring interactivity exist [17, 18, 29, 32, 34]. Measures for perceived interactivity used in this research were developed by Liu [17], Wu [34], and Song and Zinkhan [29] and they are explained subsequently in details.

Author Liu [17], developed an instrument based on conceptualization of interactive communication as a "communication that offers individuals active control and allows them to communicate both reciprocally and synchronously". Measure includes twelve items within 3 dimensions (Figure 1) of interactivity: active control, two-way communication, and synchronicity. Special care was taken to secure

that scale does not contain any attitudinal or behavioral intentions.

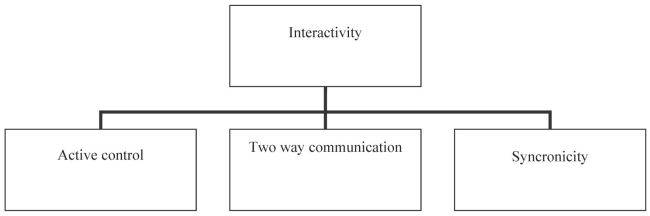



***Figure 1.*** *Liu dimensions of interactivity*

Wu [34] instrument measures interactivity using three dimensions (Figure 2) of website perceived intractivity: perceived control, perceived response, and perceived personalisation. Perceived control reflects users ability and confidence in performing activities and it assumes aspects of website navigation, the pace or rhythm of the interaction, and the content being accessed. Perceived response represents users perception of how the interactive system reacts to his/her inputs. Those responses could be from the site owner, from the navigation cues and signs, and from the real people online. Perceived personalisation is related to the extent to which users perceived appropriateness of the responses of his partner in communication as personally relevant to his communicative behaviours. Perceived personalization is analyzed through website as if it **is a** person, as if it wants to know visitor, and as if it understands the user. Instrument uses 9 statemens, 3 statemenst for each dimension.

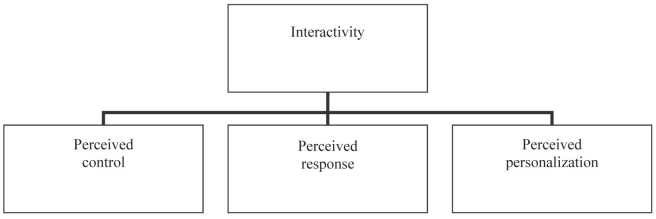



***Figure 2.*** *Wu dimensions of interactivity*

Authors Song and Zinkhan [29] using telepresence theory and interactivity theory developed instrument for measuring interactivity as a combination of Wu [32], McMillan and Hwang [18], and Liu [17] instruments. Instrument uses following dimensions (Figure 3) for measuring perceived interactivity: perceived communication, perceived control, and perceived responsiveness. Instrument has 21 items in total. Nine items are used for perceived control, whereas six items are used for perceived response and perceived communication.

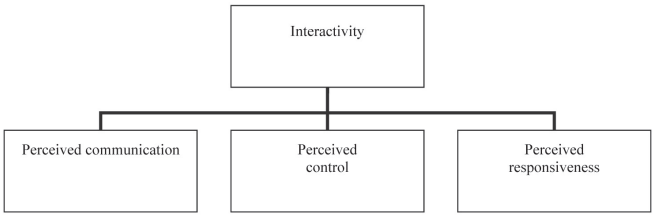



***Figure 3.*** *Song and Zinkam dimensions of interactivity*

## RESEARCH METHODOLOGY

### Participants

Participants were 120 undergraduate students selected in the pretest among the students of the freshman year at the School of Electrical Engineering of Applied Studies in Belgrade. Total number of 120 participants was divided into two groups using the principle of random selection. The equal number of males and females participated in the study. Before conducting analysis some cases were excluded from the analysis. Excluded cases had missing values and values out of the specified range. After exclusion, the number of cases employed in further analysis was 99. Table 1 shows age structure.

***Table 1.*** *Age structure of participants*

| Age | Participants (%) |
|---|---|
| Up to 20 | 72,6% |
| 21-25 | 23,9% |
| 26-30 | 2% |
| 31-40 | 1% |
| Over 40 | 0,5% |

Tables 2 and 3 show participants experience measured by the number of years someone is using the internet and hours per week spent on the internet.

***Table 2.*** *Number of years participant is using the internet*

| Years of use | Participants (%) |
|---|---|
| Less than 2 | 1% |
| 2-4 | 5% |
| 5-6 | 23,4% |
| More than 6 | 70,6% |

***Table 3.*** *Hours per week participant spent on the internet*

| Hours per week | Participants (%) |
|---|---|
| Less than 5 | 5,1% |
| 6-20 | 24,9% |
| 21-40 | 41,6% |
| More than 40 | 28,4% |

It could be concluded from the previous tables that participants were males and females of different age groups, who use the internet from 1 to over 6 years and spend from 1 to more than 40 hours per week surfing the web.

### Procedure

We conducted laboratory experiment with the aim to test selected measures of interactivity (Liu, Wu, and Song and Zinkhan) on two variants of a fictious website (low and high interactivity). Participants were recruited through pretest, using pretest questionnaire. The aim of the pretest was to identify participants as internet users with experince, and their prefered content on the internet. Prefered content was important for the development of stimuli for main research - fictious websites.

Main research was conducted in laboratory settings. Participants were randomly asigned to experimental conditions. Then, participants were informed about the study and their task. Their task was to browse website and search for the content of interest and if they found what they want to apply for training, job or internship. Participants had 30 minutes to complete the task. After they had completed the task, they received a questionnaire to fill out. It took them aproximately 20 minutes to complete the questionnaire.

### Research stimuli

Fictious web site with two versions (high interactivity and low interactivity) was developed as a research stimulus [1, 12, 13, 16, 17, 18, 29, 34]. Website content was determined during the pretest, as a most relevant content from the respondents point of view. Most relevant content for a majority of respondents were informations about trainings, internships and jobs. According to pretest, two fictious portals for trainings, internships and jobs were developed.

Websites had the same content, and the difference between websites was a level of interactivity. Interactivity manipulation was created according to recommendations presented in various researches [13, 15, 24, 28]. Websites were created using well known content management system WordPress, using template *Medicine* (version 1.0.3). High interactivity version had more interactive features than low interactivity version. Those interactive features that were integrated into the high interactivity website are: option to recommend the site to friends, option to apply for training/internship/job online, site map, e-mail hot link, on-line chat room, dynamic menu, site search, tagging, option to make a comment on the offer, sharing content via social media sites, and newsletter registration.

Low interactivity version has the following structure: Home page, Trainings section, Internships section and Jobs section. Home page contains posts with short description of actual offer from all sections of the site, job offer, internships and trainings respectively, with the link to landing page and they are organized in reverse chronological order. Beside posts, home page contains instruction with explanation of the research purpose and participant task. Disclaimer which explains the site intention is located in the footer of the site. Add-ins, positioned above the footer, contain calendar and categories offers. Other website sections and pages within sections (jobs, trainings and internships) have the same structure as home page. In the central part are posts from one of the aforementioned categories organized in reverse chronological order. Auxiliary block contains instructions. Header has the same structure as in home page. Low interactivity website page layout is shown in Figure 4.

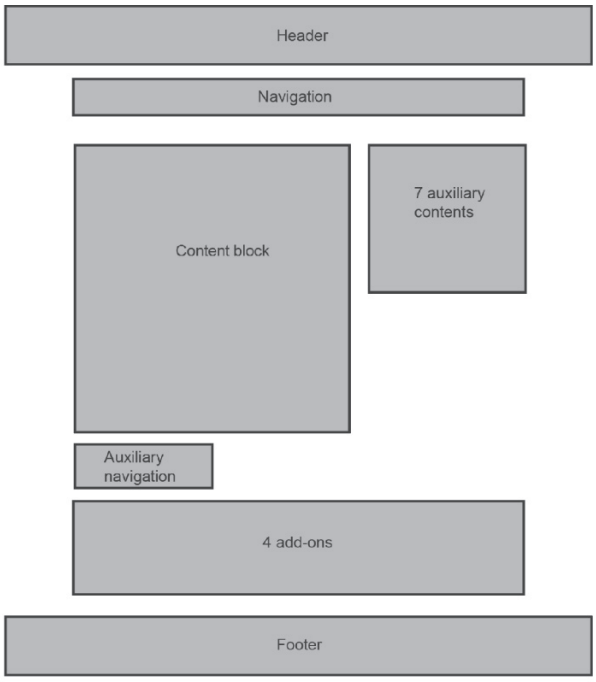



***Figure 4.*** *Layout of low interactivity website page*

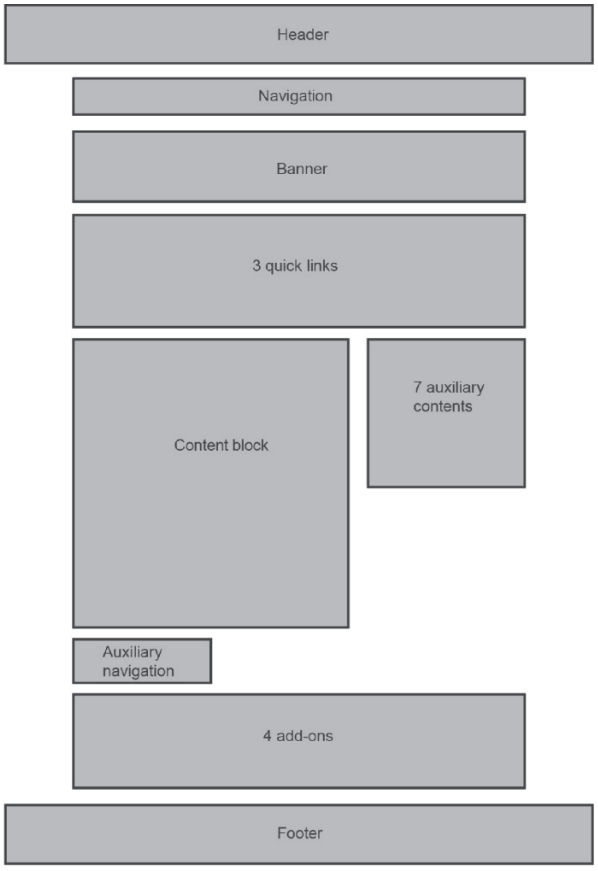



***Figure 5.*** *Layout of high interactivity website page*

High interactivity version has the following structure: Home page, Trainings section, Internships section, Jobs section, and 24hrs support. Home page contains posts like low interactivity version. Above the post block are located quick links and a banner block. Page contains 24hrs support, a registration to mailing list, Facebook social plugin, comments block, and the most wanted jobs and tags. Footer is the same as on low interactivity version. Add-ins are above footer. Section pages (jobs, trainings and internships) have the same structure as the home page, except for the shortcut links. There are single offer pages (shown on the left hand side) and pages 24hrs support which do not contain shortcut links and help navigation (displayed on the right hand side). Offer landing page opens from the navigation bar or by selecting offer from certain category. At the bottom of the offer page, users are provided with the possibility to register for selected offer using web form and to subscribe to a mailing list. Beneath the subscription form recommended offer banners and commenting fields are located. Sharing offers via social networks, Gmail and mail of user's client are offered below the picture that describes the offer and below subscription to a mailing list. Figure 5 shows the layout of high interactivity site.

## RESULTS

Total number of cases used in the analysis was 99 (51 low interactivity website, 48 high interactivity website). First, we assessed all instruments for reliability using coefficient alpha (Cronbach's alpha). Cronbach's alpha shoud have the value of 0.7 or more. Cronbach's alpha for Liu instrument was 0.943, for Wu instrument was 0.903, and for Song and Zinkam instrument was 0.876. Since all scores were larger than 0.7, therefore the instruments have high reliability. Then we calculated Cronbach's alpha for all instrument dimensions. Values are presented in Table 4, Table 5 and Table 6 respectively for Liu, Wu, and Song and Zinkam instrument. All values are equal or exceed 0.7.

***Table 4.*** *Cronbach's alpha for Liu measure dimensions*

| Dimension | Alpha |
|---|---|
| Active control | 0.7 |
| Two-way communication | 0.8 |
| Syncronicity | 0.78 |

***Table 5.*** *Cronbach's alpha for Wu measure dimensions*

| Dimension | Alpha |
|---|---|
| Perceived control | 0.81 |
| Perceived responsiveness | 0.83 |
| Perceived personalization | 0.7 |

***Table 6.*** *Cronbach's alpha for Song & Zinkan measure dimensions*

| Dimension | Alpha |
|---|---|
| Perceived communication | 0.83 |
| Perceived control | 0.76 |
| Perceived responsiveness | 0.7 |

Next we compared Liu, Wu, and Song and Zinkan measures of interactivity for participant of low interactivity versus high interactivity website. Participants show statisticaly significant ($p<0.05$) difference for all instrument for high interactivity website relative to low interactivity website (Table 7). Figure 6 graphically shows for all models the difference between mean values for high interactivity website relative to low interactivity website. The highest mean value obtained Song and Zinkan model, and Wu model obtained the lowest.

***Table 7.*** *T-test results for different interactivity measures*

| | Website variant | M | SD | t | p |
|---|---|---|---|---|---|
| **Liu** | High interactivity | 5.097 | 0.536 | 8.039 | .000 |
| | Low interactivity | 4.264 | 0.492 | | |
| **Wu** | High interactivity | 4.956 | 0.647 | 7.038 | .000 |
| | Low interactivity | 3.970 | 0.747 | | |
| **Song & Zinkan** | High interactivity | 5.187 | 0.610 | 5.636 | .000 |
| | Low interactivity | 4.485 | 0.629 | | |

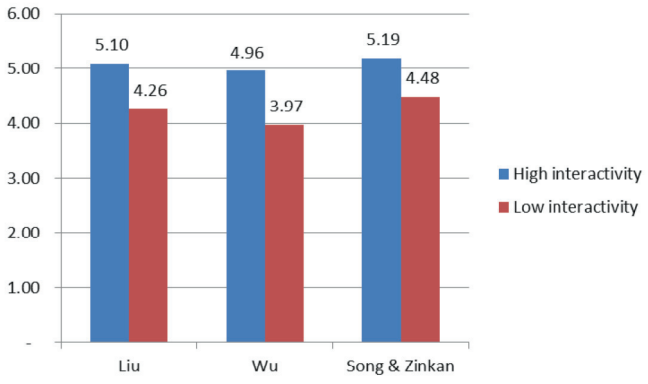



***Figure 6.*** *Graphical representations of mean data for interactivity measures*

Then we analyzed each particular dimension of the model between participants who used low interactivity and high interactivity website. Table 8, Table 9, and Table 10 show results for these models. All dimension values between participants of low and high interactivity website were statistically significant ($p<0.05$). As tables show, Liu dimension *Active control*, Wu dimension *Perceived control*, and Song and Zinkan dimension *Active control*, obtained higher mean value. These results suggest that there is a consistency of control dimension between different models. Liu *Two-way communication* dimension and Song & Zinkan dimension *Perceived communication* obtained the lowest mean values confirming that partial consistency between models exists. Other dimensions are not comparable.

***Table 8.*** *T-test results for dimensions of Liu model*

| Liu model | | | | | |
|---|---|---|---|---|---|
| | Website version | M | SD | t | p |
| **Active control** | High interactivity | 5.451 | 0.767 | 3.086 | .003 |
| | Low interactivity | 4.854 | 1.133 | | |
| **Two-way communication** | High interactivity | 4.618 | 0.973 | 6.514 | .000 |
| | Low interactivity | 3.281 | 1.069 | | |
| **Synchronicity** | High interactivity | 5.388 | 0.872 | 2.286 | .024 |
| | High interactivity | 5.451 | 0.767 | | |

***Table 9.*** *T-test results for dimensions of Wu model*

| Wu model | | | | | |
|---|---|---|---|---|---|
| | Website version | M | SD | t | p |
| **Perceived control** | High interactivity | 6.157 | 0.731 | 3.401 | .001 |
| | Low interactivity | 5.431 | 1.326 | | |
| **Perceived responsiveness** | High interactivity | 4.500 | 1.034 | 3.815 | .000 |
| | Low interactivity | 3.688 | 1.085 | | |
| **Perceived personalization** | High interactivity | 4.327 | 1.344 | 4.894 | .000 |
| | Low interactivity | 3.028 | 1.293 | | |

***Table 10.*** *T-test results for dimensions of Song & Zinkan model*

| Song & Zinkan model | | | | | |
|---|---|---|---|---|---|
| | Website version | M | SD | t | p |
| **Perceived communication** | High interactivity | 4.510 | 1.074 | 6.306 | .000 |
| | Low interactivity | 3.097 | 1.155 | | |
| **Perceived control** | High interactivity | 5.885 | 0.685 | 3.861 | .000 |
| | Low interactivity | 5.162 | 1.137 | | |
| **Perceived responsiveness** | High interactivity | 5.124 | 0.745 | 2.169 | .033 |
| | Low interactivity | 3.028 | 1.293 | | |

Graphical representations for all model dimensions between participants who used low interactivity and high interactivity website are shown in Figure 7, Figure 8, and Figure 9.

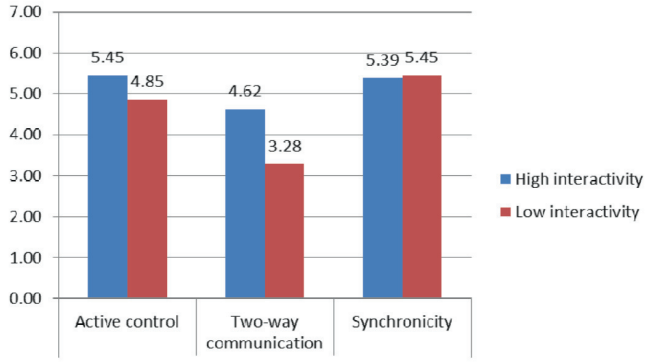



***Figure 7.*** *Graphical representations of mean data for Lui model dimensions*

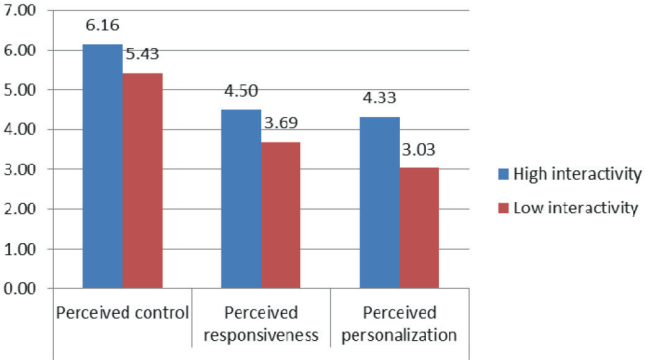



***Figure 8.*** *Graphical representations of mean data for Wu model dimensions*

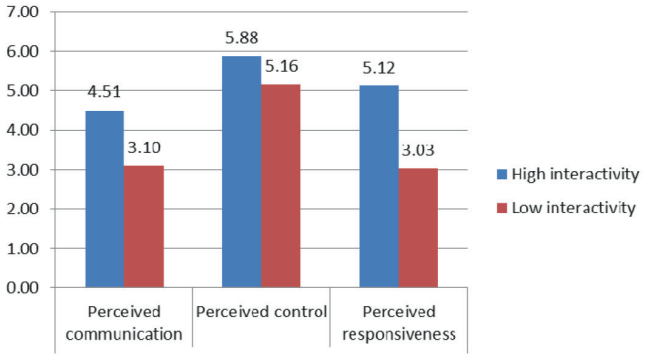



***Figure 9.*** *Graphical representations of mean data for Song & Zinkan model dimensions*

We also analyzed for selected models every particular question for low interactivity website and high interactivity website. First, we analyzed Liu model, then Wu model, and Song & Zinkan model. Results are shown in Table 11, Table 12, and Table 13.

***Table 11.*** *Responses of the participants of low interactivity and high interactivity site for Liu model*

| | Website version | M | SD |
|---|---|---|---|
| **Active control** | | | |
| I felt that I had a lot of control over my visiting experiences at this website. | High interactivity | 5.098 | 1.315 |
| | Low interactivity | 4.417 | 1.674 |
| While I was on the website, I could choose freely what I wanted to see. | High interactivity | 6.431 | 1.005 |
| | Low interactivity | 5.875 | 1.684 |
| While surfing the website, my actions decided the kind of experiences I got. | High interactivity | 5.471 | 1.332 |
| | Low interactivity | 4.708 | 1.429 |
| While surfing the website, I had absolutely no control over what I can do on the site. | High interactivity | 4.804 | 2.050 |
| | Low interactivity | 4.417 | 1.855 |
| **Two-way communication** | | | |
| This website facilitates two-way communication between the visitors and the site. | High interactivity | 4.745 | 1.339 |
| | Low interactivity | 3.750 | 1.995 |
| The website makes me feel it wants to listen to its visitors. | High interactivity | 4.294 | 1.932 |
| | Low interactivity | 2.333 | 1.521 |
| Site created the feeling that it wants to listen to its users | High interactivity | 4.941 | 1.462 |
| | Low interactivity | 3.771 | 1.547 |
| The website gives visitors the opportunity to talk back. | High interactivity | 5.196 | 1.096 |
| | Low interactivity | 3.521 | 1.624 |
| It is difficult to offer feedback to the website. | High interactivity | 4.373 | 1.296 |
| | Low interactivity | 3.333 | 1.492 |
| The website does not at all encourage visitors to talk back. | High interactivity | 4.157 | 1.869 |
| | Low interactivity | 2.979 | 1.828 |
| **Synchronicity** | | | |
| The website processed my input very quickly. | High interactivity | 5.157 | 1.155 |
| | Low interactivity | 4.667 | 1.404 |
| Getting information from the website is very fast. | High interactivity | 5.667 | 1.194 |
| | Low interactivity | 5.354 | 1.280 |
| I was able to obtain the information I wanted without any delay. | High interactivity | 5.549 | 1.137 |
| | Low interactivity | 5.083 | 1.350 |
| When I clicked on the links, I felt I was getting instantaneous information. | High interactivity | 5.098 | 1.664 |
| | Low interactivity | 4.792 | 1.320 |
| The website was very slow in responding to my requests. | High interactivity | 5.471 | 1.759 |
| | Low interactivity | 4.958 | 1.924 |

***Table 12.*** *Responses of the participants of low interactivity and high interactivity site for Wu model*

| | Website version | M | SD |
|---|---|---|---|
| **Perceived control** | | | |
| I was in control of my navigation through this website. | High interactivity | 6.275 | 1.185 |
| | Low interactivity | 5.396 | 1.865 |
| I had some control over the content of this website that I wanted to see. | High interactivity | 6.275 | 1.060 |
| | Low interactivity | 5.375 | 1.721 |
| I was in total control over the pace of my visit to this website. | High interactivity | 5.922 | 1.246 |
| | Low interactivity | 5.521 | 1.304 |
| **Perceived responsiveness** | | | |
| I could communicate with the company directly for further questions about the company or its products if I wanted to. | High interactivity | 4.294 | 1.932 |
| | Low interactivity | 2.333 | 1.521 |
| I could communicate in real time with other customers who shared my interest in this website. | High interactivity | 4.333 | 1.873 |
| | Low interactivity | 2.771 | 1.801 |
| The site had the ability to respond to my specific questions quickly and efficiently | High interactivity | 5.196 | 1.096 |
| | Low interactivity | 3.521 | 1.624 |
| **Perceived personalization** | | | |
| I perceived the website to be sensitive to my needs for product information. | High interactivity | 3.804 | 1.470 |
| | Low interactivity | 3.854 | 1.502 |
| I felt I just had a personal conversation with a sociable, knowledgeable and warm representative from the company. | High interactivity | 4.353 | 1.585 |
| | Low interactivity | 3.979 | 1.657 |
| The Web site was like talking back to me while I clicked through the website. | High interactivity | 5.098 | 1.664 |
| | Low interactivity | 4.792 | 1.320 |

***Table 13.*** *Responses of the participants of low interactivity and high interactivity site for Song & Zinkan model*

| | Website version | M | SD |
|---|---|---|---|
| **Perceived communication** | | | |
| This Web site facilitates two-way communication. | High interactivity | 4.745 | 1.339 |
| | Low interactivity | 3.750 | 1.995 |
| The Web site gives me the opportunity to talk back. | High interactivity | 4.294 | 1.932 |
| | Low interactivity | 2.333 | 1.521 |
| The Web site facilitates concurrent communication. | High interactivity | 4.333 | 1.873 |
| | Low interactivity | 2.771 | 1.801 |
| The Web site enables conversation. | High interactivity | 4.333 | 1.956 |
| | Low interactivity | 3.229 | 1.666 |
| The site is effective in gathering visitors' feedback. | High interactivity | 5.196 | 1.096 |
| | Low interactivity | 3.521 | 1.624 |
| The Web site does not encourage visitors to talk back. | High interactivity | 4.157 | 1.869 |
| | Low interactivity | 2.979 | 1.828 |
| **Perceived control** | | | |
| While I was on the site, I was always aware where I was. | High interactivity | 6.177 | 1.260 |
| | Low interactivity | 5.313 | 1.959 |
| While I was on the site, I always knew where I was going. | High interactivity | 6.098 | 1.082 |
| | Low interactivity | 5.208 | 1.821 |
| While I was on the site, I could choose freely what I wanted to see. | High interactivity | 6.431 | 1.005 |
| | Low interactivity | 5.875 | 1.684 |
| While surfing the site, my actions decided the kind of experiences I got. | High interactivity | 5.471 | 1.332 |
| | Low interactivity | 4.708 | 1.429 |
| While I was on the site, I was always able to go where I thought I was going. | High interactivity | 6.059 | 1.139 |
| | Low interactivity | 5.250 | 1.644 |
| I was delighted to be able to choose which link and when to click. | High interactivity | 6.137 | 1.510 |
| | Low interactivity | 5.625 | 1.482 |
| While surfing the site, I had absolutely no control over what I could do on the site. | High interactivity | 4.804 | 2.050 |
| | Low interactivity | 4.417 | 1.855 |
| The Web site is not manageable. | High interactivity | 5.902 | 1.285 |
| | Low interactivity | 4.896 | 1.741 |
| I feel that I have a great deal of control over my visiting experience at this site. | High interactivity | 5,098 | 1,315 |
| | Low interactivity | 4,417 | 1,674 |
| **Perceived responsiveness** | | | |
| The Web site processed my input very quickly. | High interactivity | 5.157 | 1.155 |
| | Low interactivity | 4.667 | 1.404 |
| Getting information from the Web site is very fast. | High interactivity | 5.667 | 1.194 |
| | Low interactivity | 5.354 | 1.280 |
| I was able to obtain the information I want without any delay. | High interactivity | 5.549 | 1.137 |
| | Low interactivity | 5.083 | 1.350 |
| When I clicked on the links, I felt I was getting instantaneous information. | High interactivity | 5.098 | 1.664 |
| | Low interactivity | 4.792 | 1.320 |
| The Web site was very slow in responding to my request. | High interactivity | 5.471 | 1.759 |
| | Low interactivity | 4.958 | 1.924 |
| The Web site answers my question immediately. | High interactivity | 3.804 | 1.470 |
| | Low interactivity | 3.854 | 1.502 |

## Conclusion

As we emphasized in previous sections, interactivity is a multidimensional and complex concept. Additional difficulty is a fact that interactivity can be actual and perceptual. Actual interactivity can be operationalized by using different interactive fea-

tures. Recommendations are numerous, but sometimes confronting and out of date. Some researchers argue that using a lot of interactivity features could harm interactivity perception [3, 16]. Various studies confirmed that more features mean higher perceptual interactivity score. But large number of studies was conducted ten and more years ago. In the meantime, new technologies changed the face of interactive systems. We want to explore whether measures of interactivity can cope with all those changes. In our case, we manipulated several interactivity features, both traditional (site map, on-line chat room or site search) and new one (online application, sharing content via social media or tagging). Websites in the experiment were differentiated only in terms of interactivity. Results confirmed that all selected measures of perceptual interactivity (Liu, Wu, and Song and Zinkan) determined statistically significant difference between participant who used low interactivity and participants who used high interactivity website. This is important for practice because interactivity can be measured with different models regardless of implemented features.

Sometimes practitioners without intention to favor, put emphasize, for example, on two way interaction or personalization. In this way, some dimension could be neglected. Interactivity is multidimensional, and all dimensions need equal attention. As we said, several measures of interactivity we use in our research could assess interactivity very well, but the question remains whether they are good enough to assess every dimension (for example perceived communication, perceived control or perceived responsiveness). In our study, we analyzed whether different values of various interactivity dimensions estimated for high and low interactivity websites are statistically significant. Our study found that almost all models and their subsequent dimensions are good predictors of interactivity. Liu model dimensions, namely active control and two-way communication were statistically significant between low and high interactivity websites. Mean values for those two dimensions, as we expected, were higher for high interactivity website. But for synchronicity dimension, mean value was lower. All Wu model dimensions values were higher for high interactivity website. Differences between dimension values for high interactivity website and low interactivity website were statistically significant. And for Song and Zinkan model, all dimension values for high and low interactivity website were statistically significant, and with higher mean values for high interactivity website.

Our research has several limitations. First, it is relatively small sample size (n=99). Further, researh could use larger samples. Second limitation is the stimulus. All interactivity features were implemented on one website. This could have impact on the complexity of browsing and searching for relevant topics. Future research could include diferent web sites presented with features, for example (low, midle, and high interactivity). Another way of thinking coud be diferent websites in terms of different interactivity dimensions. Future research could try to identify a linkage between interactivity level and the effects of interactivity.

## References


[1] Albert TC, Goes PB, and Gupta A (2004). GIST: a model for design and management of content and interactivity of customer-centric web sites. *MIS Quarterly*, 28 (2), 161-182.

[2] Barwise P and Farley JU (2005) The state of interactive marketing in seven countries: Interactive marketing comes of age. *Journal of interactive marketing*, 19 (3), 67–80.

[3] Bezjian-Avery A Calder B and Iacobucci D (1998) New media interactive advertising vs. traditional advertising. *Journal of advertising research*, 38, 23-32.

[4] Bird D (2007) Commonsense Direct & Digital Marketing (5th ed.). Kogan Page Ltd, United Kingdom.

[5] Charlesworth A (2014) Digital marketing. 2nd edition. Taylor & Francis Group, Routlege. New York. USA.

[6] Coyle JR. and Thorson E (2001) The Effects of Progressive Levels of Interactivity and Vividness in Web Marketing Sites. Journal of advertising, 30 (3), 65–77.

[7] Downes EJ and McMillan SJ (2000) Defining interactivity: a qualitative identification of key dimensions. *New Media and Society*, 2 (2), 157–179.

[8] Econsultancy (2016) Marketing budgets 2016. http://demand.eloqua.com/LP=6546

[9] Ha L and James LE (1998) Interactivity reexamined: A baseline analysis of early business Web sites. Journal of Broadcasting & Electronic Media, 42 (4), 456–474.

[10] Halligan B and Shah D (2009) *Inbound Marketing: Get Found Using Google, Social Media and Blogs*. John Wiley & Sons.

[11] Jensen JJ (1998) Interactivity: Tracking a New Concept in Media and Communication Studies. *Nordicom Review*, 19 (1), 185-204.

[12] Jiang Z Chan J Tan BCY and Chua WS (2010) Effects of Interactivity on Website Involvement and Purchase Intention. *Journal of the Association for Information Systems*, 11(1), 34–59.

[13] Johnson G Bruner G and Kumar A (2006) Interactivity and Its Facets Revisited: Theory and Empirical Test, *Journal of Advertising, 35*(4), 35-52.

[14] Kiousis S (2002) Interactivity: a concept explication. *New Media & Society*, 4 (3), 355-383.

[15] Leone C (2016) How Much Should You Budget For Marketing In 2017? https://www.webstrategiesinc.com/blog/how-much-budget-for-online-marketing-in-2014

[16] Liu YP and Shrum LJ (2002) What is interactivity and is it always such a good thing? Implications of definition, person and situation for the influence of interactivity on advertising effectiveness. *Journal of Advertising*, 31 (4), 53–64.

[17] Liu YP (2003) Developing a scale to measure the interactivity of web sites. *Journal of Advertising Research*, 43, 207-216.

[18] McMillan SJ and Hwang JS (2002) Measures of Perceived Interactivity: An Exploration of the Role of Direction of Communication User Control and Time in Shaping Perceptions of Interactivity. Journal of advertising, 31 (3), 29–42.

[19] McMillan SJ (2002) A four-part model of cyber-interactivity. *New Media & Society*, SAGE Publications, 4 (2), 271–291.

[20] Mohammed RA Fisher RJ Jaworski BJ and Cahill AM (2002) Internet Marketing - Building Advantage in a Networked Ecconomy. McGraw-Hill Irwin, New York. NY.

[21] Rafaeli S (1988) Interactivity: From New Media to Communication. In Advancing Communication Science: Merging Mass and Interpersonal Processes. (Hawkins RP Wiemann JM and Pingree S) pp 110-134, Sage, Beverly Hills. CA.

[22] Rafaeli S and Sudweeks F (1997) Networked Interactivity. *Journal of Computer-Mediated Communication*, 2(4)

[23] Rafaeli S and Yaron A (2009) Assessing Interactivity in Computer-Mediated Research In The Oxford Handbook of Internet Psychology (Joinson AN McKenna KYA Postmes T and Rieps UD), pp. 71-88, Oxford University Press.

[24] Reed J (2012) Get Up to Speed with Online Marketing, Upper Saddle River, New Jersey, FT Press

[25] Rogers EM (1986) Communication Technology. The New Media in Society. New York.

[26] Ryan D (2014) Understanding digital marketing – Marketing strategies for engaging the digital generation. Third edition. Kogan Page Ltd.

[27] Shimp T (2006) Advertising Promotion and Other Aspects of Integrated Marketing Communications. 7th ed. Cengage Learning.

[28] Silverstein B (2002) Business-to-Business Internet Marketing. Fourth Edition, Maximum Press.

[29] Song JH and Zinkhan GM (2008) Determinants of Perceived Web Site Interactivity. *Journal of marketing*, 72 (2), 99-133.

[30] Steuer J (1992) Defining Virtual Reality: Dimensions Determining Telepresence, *Journal of communication*, 42 (4), 73–93.

[31] Strauss J and Frost R (2012). E-Marketing, 6th edition, Pearson Education – Prentice Hall Inc, Upper Saddle River, New Jersey, USA.

[32] Wu G. (1999) Perceived Interactivity and Attitude toward Website. Annual Conference of American Academy of Advertising,

Albuquerque. New Mexico, pp 254-262.

[33] Wu G (2005) The mediating role of perceived interactivity in the effect of actual interactivity on attitude toward the website. *Journal of Interactive Advertising*, 5 (2), 29–39.

[34] Wu G (2006) Conceptualizing and Measuring the Perceived Interactivity of Websites. *Journal of Current Issues & Research in Advertising,* 28 (1), 87–104.

[35] Yadav MS and Varadarajan R (2005) Interactivity in the electronic marketplace: an exposition of the concept and implications for research. *Journal of the Academy of Marketing Science*, 33 (4), 585-603.

[36] Zack MH (1993) Interactivity and Communication Mode Choice in Ongoing Management Groups. *Information Systems Research*, 4 (3), 207–239.

Submitted: May 9, 2017.
Accepted: June 13, 2017.

## ABOUT THE AUTHORS

**Velimir Štavljanin** is an associate professor of marketing management and multimedia communications and a senior researcher in Laboratory for Multimedia Communications at the University of Belgrade, Faculty of organizational sciences, Serbia. His areas of interests include product and brand management, digital marketing, CRM, multimedia and HCI. Štavljanin has a PhD in marketing from the University of Belgrade, Faculty of organizational sciences.

**Milica Jevremović** was born in 1980. In 2005, she graduated from the Faculty of Organizational Sciences, University of Belgrade, course Management, defending the thesis "Market research". Mrs. Jevremović was awarded with the PhD Title by defending the PhD dissertation "Research of the interactivity in digital marketing strategies", course marketing in June 2016 at the aforementioned Faculty. Following graduation in 2006, Mrs. Jevremovic established employment relationship with the School of Electrical Engineering and IT of Applied Studies (VISER) in Belgrade, as an expert associate. As of 2009, Mrs. Jevremovic is acting in the capacity of a fellow assistant at the VISER. Since 2016, Mrs. Jevremovic is engaged in the capacity of a lecturer at the same Institution.

# Instructions for Authors

The *Journal of Information Technology and Application (JITA)* publishes quality, original papers that contribute to the methodology of IT research as well as good examples of practical applications.

Authors are advised that adherence to the Instructions to Authors will help speed up the refereeing and production stages for most papers.

- Language and presentation
- Length of submissions
- Submission
- Contact details/biographies
- Title of the paper
- Abstract and keywords
- Figures and tables
- Sections
- Footnotes
- Special characters
- Spelling
- References
- Proofs
- PDF offprint
- Copyright and permissions
- Final material
- Correspondence
- Publication ethics

## Language and presentation

Manuscripts should be written in English. All authors should obtain assistance in the editing of their papers for correct spelling and use of English grammar. Manuscripts should have double spacing, with ample margins and pages should be numbered consecutively. The Editors reserve the right to make changes that may clarify or condense papers where this is considered desirable.

## Length of submissions

Papers should not normally exceed 12 Journal pages (about 8000 words). However, in certain circumstances (e.g., review papers) longer papers will be published.

## Submission

Manuscripts must be submitted through the JITA online submission system.

Please read the instructions carefully before submitting your manuscript and ensure the main article files do not contain any author identifiable information.

Although PDF is acceptable for initial submission original source (i.e. MS Word) files will be required for typesetting etc.

## Contact details/biographies

A separate file containing the names and addresses of the authors, and the name and full contact details (full postal address, telephone, fax and e-mail) of the author to whom correspondence is to be directed should be uploaded at the time of submission (you should select Contact details/Biographies as the file type). This file is not shown to reviewers. This file should also contain short biographies for each author (50 words maximum each) which will appear at the end of their paper.

The authors' names and addresses must not appear in the body of the manuscript, to preserve anonymity. Manuscripts containing author details of any kind will be returned for correction.

## Title of the paper

The title of the paper should not be longer than 16 words.

## Abstract and keywords

The first page of the manuscript should contain a summary of not more than 200 words. This should be self-contained and understandable by the general reader outside the context of the full paper. You should also add 3 to 6 keywords.

## Figures and tables

Figures which contain only textual rather than diagrammatic information should be designated Tables. Figures and tables should be numbered consecutively as they appear in the text. All figures and tables should have a caption.

## Sections

Sections and subsections should be clearly differentiated but should not be numbered.

## Footnotes

Papers must be written without the use of footnotes.

## Special characters

Mathematical expressions and Greek or other symbols should be written clearly with ample spac-

ing. Any unusual characters should be indicated on a separate sheet.

## Spelling

Spelling must be consistent with the Concise Oxford Dictionary.

## References

References in the text are indicated by the number in square brackets. If a referenced paper has three or more authors the reference should always appear as the first author followed by et al. References are listed alphabetically. All document types, both printed and electronic, are in the same list. References to the same author are listed chronologically, with the oldest on top. Journal titles should not be abbreviated.

**Journal**

Avramović ZŽ (1995) Method for evaluating the strength of retarding steps on a marshalling yard hump. *European Journal of Operational Research*, 85(1), 504–514.

**Book**

Walsham G (1993) Interpreting Information Systems in Organizations. Wiley, Chichester.

**Contributed volume**

Huberman AM and Miles MB (1994) Data Management and analysis methods. In Handbook of Qualitative Research (Denzin NK and Linkoln YS, Eds), pp 428-444, Sage, Thousand Oaks, California.

**Conference Paper**

Huberman AM and Miles MB (1994) Data Management and analysis methods. In Handbook of Qualitative Research (Denzin NK and Linkoln YS, Eds), pp 428-444, Sage, Thousand Oaks, California.

**Unpublished reports/theses**

Nandhakumar JJ (1993) The practice of executive information systems development: and in-depth case study. PhD Thesis, Department of Engineering, University of Cambridge.

## Proofs

Proofs of papers will be sent to authors for checking. Alterations to diagrams should be avoided where possible. It will not be possible to accept major textual changes at this stage. Proofs must be returned to the publishers within 48 hours of receipt by fax, first-class post, airmail or courier. Failure to return the proof will result in the paper being delayed.

## PDF offprint

Corresponding authors will receive a PDF of their article. This PDF offprint is provided for personal use. It is the responsibility of the corresponding author to pass the PDF offprint onto co-authors (if relevant) and ensure that they are aware of the conditions pertaining to its use.

The PDF must not be placed on a publicly-available website for general viewing, or otherwise distributed without seeking our permission, as this would contravene our copyright policy and potentially damage the journal's circulation. Please visit **http://www.apeiron-journals.com/JITA/authors/rights_and_permissions.html** to see our latest copyright policy.

## Copyright and permissions

The copyright of all material published in the Journal is held by Paneuropean University APEIRON. The author must complete and return the copyright form enclosed with the proofs.

Authors may submit papers which have been published elsewhere in a foreign language, provided permission has been obtained from the original publisher before submission.

Authors wishing to use material previously published in JITA should consult the publisher.

## Final material

All final material must be submitted electronically in its original application format (MS Word is preferred). The file must correspond exactly to the final version of the manuscript.

## Correspondence

Business correspondence and enquiries relating to advertising, subscriptions, back numbers or reprints should be addressed to the relevant person at:

Paneuropean University APEIRON
Journal JITA
Pere Krece 13, P.O.Box 51
78102 Banja Luka
Bosnia and Hercegovina / RS

## Publication ethics

We take an active interest in issues and developments relating to publication ethics, such as plagiarism, falsification of data, fabrication of results and other areas of ethical misconduct. Please note that submitted manuscripts may be subject to checks using the corresponding service, in order to detect instances of overlapping and similar text.

# JITA

PUBLISHER
Paneuropean University APEIRON,
College of Information Technology
Banja Luka, Republic of Srpska, B&H
**www.apeiron-uni.eu**
**Darko Uremović**, Person Responsible for the Publisher
**Aleksandra Vidović**, Editor of University Publications



ISSN 2232-9625

Journal of Information Technology and Applications VOLUME 7 NUMBER 1 JUNE 2017