Infocommunications Journal

HTE 75 1949–2024

A PUBLICATION OF THE SCIENTIFIC ASSOCIATION FOR INFOCOMMUNICATIONS (HTE)

ISSN 2061-2079

Special Issue

Joint Special Issue on Cognitive Infocommunications and Cognitive Aspects of Virtual Reality

Technically Co-Sponsored by

IEEE ComSoc™ IEEE Communications Society

hte

IEEE HUNGARY SECTION

Editorial Board

Indexing information

Infocommunications Journal is covered by Inspec, Compendex and Scopus.
Infocommunications Journal is also included in the Thomson Reuters – Web of ScienceTM Core Collection, Emerging Sources Citation Index (ESCI)

Infocommunications Journal

Technically co-sponsored by IEEE Communications Society and IEEE Hungary Section

Supporters

FERENC VÁGUJHELYI – president, Scientific Association for Infocommunications (HTE)

The publication was produced with the support of the Hungarian Academy of Sciences and the NMHH

National Media and Infocommunications Authority • Hungary

Editorial Office (Subscription and Advertisements)**:**
Scientific Association for Infocommunications
H-1051 Budapest, Bajcsy-Zsilinszky str. 12, Room: 502
Phone: +36 1 353 1027 • E-mail: info@hte.hu • Web: www.hte.hu

Articles can be sent also to the following address:
Budapest University of Technology and Economics
Department of Telecommunications and Media Informatics
Phone: +36 1 463 4189 • E-mail: pvarga@tmit.bme.hu

Subscription rates for foreign subscribers: 4 issues 10.000 HUF + postage

Publisher: PÉTER NAGY

HU ISSN 2061-2079 • Layout: PLAZMA DS • Printed by: FOM Media

Joint Special Issue on Cognitive Infocommunications and Cognitive Aspects of Virtual Reality

Péter Baranyi, Atsushi Ito, Ádám B. Csapó, Ildikó Horváth, and Tibor Guzsvinecz

We are pleased to introduce this joint special issue focusing on Cognitive Infocommunications (CogInfoCom) and Cognitive Aspects of Virtual Reality (cVR). Within this Issue a wide range of research results are presented, focusing on the complex relationship between human cognition and information and communication technologies (ICT), along with the revolutionary effects of virtual reality settings on cognitive functions.

The first paper presents a novel methodology for clustering texts based on word frequency analysis of English translations of ancient Greek texts. Leveraging the classification system of the ancient Library of Alexandria, this research offers insights into textual patterns and cognitive processes underlying literary masterpieces.

The second paper explores environmental factors shaping eighth-grade pupils' mathematical abilities. Utilizing machine learning methods, critical elements influencing mathematical proficiency are identified, highlighting the importance of holistic instructional efforts.

The third paper investigates human perception complexities through an analysis of memory colors in the CIELAB color space. This study underscores the dynamic interplay between cognition and sensory experiences in virtual environments.

The fourth paper addresses human-computer interaction interfaces in telerehabilitation for individuals with movement disabilities. Innovative algorithms offer promising avenues for enhancing gesture recognition accuracy and usability.

The fifth paper presents recent transformative changes in Cognitive Infocommunications, emphasizing the entangled nature of human-technology interactions and the co-evolution of cognition and ICT.

The sixth paper underscores the importance of web accessibility, proposing an approach for computing the accessibility score of academic web pages, promoting inclusivity in digital environments.

The seventh paper introduces an artificial cognitive and multilabel classification approach to expedite the tagging process of legal documents, facilitating efficient document classification.

The eighth paper presents a comprehensive overview of the 3D virtual library project, showcasing its transformative impact on language learning experiences for computer science students.

The ninth paper introduces the Internet of Digital Education (IoDE) model to facilitate seamless and inclusive learning experiences, leveraging modern technologies in advancing educational practices.

The tenth paper explores the cognitive aspects of virtual reality systems through an analysis of textual reviews of virtual reality games, providing insights to inform the refinement of immersive virtual environments.

The eleventh paper broadens the concept of avatar design to incorporate displays conveying interpersonal information, enhancing social interactions in digital platforms.

The twelfth paper addresses critical challenges in road safety through innovative driver assistance systems tailored to drivers' emotional needs.

The thirteenth paper introduces a system of fuzzy indicators to mitigate cognitive biases in logistics decision-making, promoting informed decision-making.

The fourteenth paper examines design principles for 3D Virtual Reality environments, offering actionable guidelines for optimizing task performance efficiency in immersive virtual environments.

We hope this special issue inspires further interdisciplinary research and innovation in these fields.

Guest Editors of the Joint Special Issue
Prof. Péter Baranyi, *Corvinus University of Budapest*
Prof. Atsushi Ito, *Chuo University*
Dr. Ádám Csapó, *Corvinus University of Budapest*
Dr. Ildikó Horváth, *Corvinus University of Budapest*
Dr. Tibor Guzsvinecz, *University of Pannonia*

Special Issue of the Infocommunication Journal

Deep Learning-Based Analysis of Ancient Greek Literary Texts in English Version: A Statistical Model Based on Word Frequency and Noise Probability for the Classification of Texts

Zoltan Gal, and Erzsébet Tóth

***Abstract*—In our paper we intend to present a methodology that we elaborated for clustering texts based on the word frequency in the English translations of selected old Greek texts. We used the classification system of the ancient Library of Alexandria, devised by the prominent Greek scholar-poet, Callimachus in the 3rd century BC., as a basis for categorizing literary masterpieces. In our content analysis, we could determine a triplet of a, b, c values for describing a power function that appropriately fits a curve determined by the word frequencies in the texts. In addition, we have discovered 16 special features of the different texts that correspond to various token categories investigated in each text, such as part of speech of the word in the context, numerals, subordinate conjunction, symbols, etc. We have developed a cognitive model in which several hundred different subtexts were utilized for supervised learning with the aim of subtext class recognition. Concerning 200 subtexts, the triplet of a, b, c values, the classes of the subtexts, and their 16-dimensional feature vectors were learnt for the Recurrent Neural Network (RNN). It turned out that the Long-Short Term Memory RNN could efficiently predict which class a chosen subtext could be categorized into without considering the interpretation of the content. The influence of the non-zero error rate of new communication services on the meaning of the transferred texts was also investigated. The impact of the noise on the classification accuracy was found to be linear, dependent on the character error rate**

***Index Terms*—deep learning, old Greek literary texts, Pinakes, automatic content analysis, text classification, Recurrent Neural Network (RNN), Long-Short Term Memory, noisy texts.**

I. Introduction

In the growing methodological pool of modern social science, content analysis is still a frequently used method [1] and one of the most significant research techniques [2, 3] applicable to quantitative and qualitative data. This method allows verbal, written, and visual data to be thoroughly investigated, systematically analyzed, and categorized. In the Internet era, the use of social media platforms is becoming inevitable for the members of the generation Cognitive Entities (CE). Content analysis can also be used for examining networked short texts [4]. There is no doubt that the careful reading and manual processing of these texts is a very time-consuming task for researchers. For this reason, automatic content analysis is applied to predict users' positive or negative preferences, and interests based on their short messages on social media. This new type of content analysis which makes possible the classification of users' feelings e.g. in the case of IMDB movie film reviews is called sentiment analysis in the field of deep learning [5]. The role of artificial intelligence and machine learning has become more and more important currently in the extraction of meaningful information and knowledge from different types of Big Data sources [14, 15].

Computer software could be used effectively at two different stages of textual content analysis: first, for storing, examining, and reporting research data; second, for automatic screening of texts, identifying and coding words and expressions [6]. An algorithmic solution could be applied in six steps: i) splitting texts into various segments; ii) determining similarities; iii) clustering; iv) cluster labeling; v) examining categories; vi) presentation of results in detail [7]. In these text analyses supervised or unsupervised automatic classification could be utilized. The researcher has to modify and prepare the texts for automatic content analysis. The preparation phase is composed of gathering textual data from social media or networks, preprocessing the texts, and defining research aims [4].

In section two a short overview of the related works is presented in the context of virtual library projects. Section three describes the analysis methodology accomplished on 37 ancient Greek texts translated into English. After highlighting the basic statistical properties of those texts, their preprocessing and neural network-based evaluation are discussed. Section four concludes and recommends the possible continuation of this research topic.

II. Related Works

The famous Great Library of Alexandria gathered all the available copies of the authentic old Greek literary texts from ancient times. In this respect, it could be regarded as a universal library representing a frequently cited symbol of human knowledge till nowadays. In its collection, the ancient texts were arranged in the alphabetical order of author names and

Z. Gal is with Faculty of Informatics, University of Debrecen, Debrecen, Hungary (e-mail: gal.zoltan@inf.unideb.hu)
E. Tóth is with Faculty of Informatics, University of Debrecen, Debrecen, Hungary (e-mail: toth.erzsebet@inf.unideb.hu).

DOI: 10.36244/ICJ.2024.5.1

classified according to their literary genres based on the Pinakes compiled by the illustrious Greek scholar, Callimachus in the 3rd century BC. In the framework of the cognitive infocommunications (CogInfoCom) research [8, 9, 10], a virtual library project was launched in 2013 to develop the three-dimensional virtual library model (3DVLM) of the ancient Library of Alexandria by collecting the relevant verbal and multimedia content about it [11]. Later this project continued to provide an overview of the greatest achievements of Callimachus by presenting the web content about his life and literary works in carefully arranged smartboards in the 3D Castle space of the MaxWhere Seminar System [12, 13].

The famous paper demonstrates that the source language of medium-length speeches in the EUROPARL corpus can often be identified through frequency counts of word n-grams, achieving an accuracy between 87.2% and 96.7% depending on the classification method. The study delves into identifying powerful positive markers and examining linguistic, cultural, and domain-related aspects. When considering all six target language versions, classification is more accurate compared to situations with only a single version available. The research also explores the diverse nature of strong markers, encompassing vocabulary, discourse structure, syntax, and contrasts between source and target languages, emphasizing the need for harmonization in terminology for improved information retrieval in parliamentary proceedings. The paper highlights the importance of further research to understand translation processes, automation of target and source language phrase retrieval, and the potential incorporation of source language considerations in EUROPARL translation models [18].

In their paper, prominent authors outline experiments in fine-grained stylometric analysis for distinguishing features among contemporaneous literary translations, both parallel and non-parallel, of works by the same author. Focusing initially on translations of Henrik Ibsen's drama "Ghosts", the study extends to explore prose translations of Anton Chekhov by different translators, such as Marian Fell and Constance Garnett. Various machine learning approaches, including Support Vector Machines and Decision Tree classifiers, are employed to identify textual features, comparing their frequencies in translations to those in reference corpora. The study successfully establishes that common word unigrams and bigrams serve as salient features for classifying translator fingerprints, achieving accuracy measurements exceeding 90%, and reveals distinctive stylistic traits between translations by William Archer and R. Farquharson Sharp of Ibsen. Cross-validation experiments suggest the effectiveness of both document-level features and n-grams in distinguishing translators, with clustering experiments showcasing differences in text clustering based on frequent words versus discriminating words identified through supervised machine learning [19].

A frequently cited study explores differences between translated and non-translated texts, acknowledging two sources of variation: interference from the source language and general effects of the translation process. The authors identify consistent effects across texts translated from the same source language through text categorization experiments, indicating a continuum between source-language-specific and general translation effects. Remarkably, classifiers based solely on function words can accurately distinguish translated from original texts, even across unrelated source languages and multiple genres. The findings suggest that both the source language and the act of translation significantly influence the characteristics of translated texts, and the observed differences align with linguistic typology, offering insights into language distances and potential applications in computational tasks, such as improving machine translation and constructing language models [20].

III. Analysis Methodology

Ancient texts, like modern texts, describe a sequence of events in chronological order, even in multiple time planes. A time plane is a story happening in a dedicated geographical area in a limited time interval. It is a crucial time structure that provides logical connections to other parts of the literary work, offering a convergent evolution of the whole story. The author of the text uses his style in the narration of the story. This fact generates the cohesion of that text, which can be used in the automatic characterization of each subtext belonging to the main text. The analysis methodology presented in the following proves a strong correlation between the patterns of part of speech frequencies and the classes of texts.

Kaggle Database contains 1/3 million most frequent English Words on the Web [16]. The relative frequency of the words versus rank is represented in Fig. 1. Rank in this context means an order of the number of words in the analyzed text. The fitting equation of the intensity curve we found to be given the following formula:

$$y(x) = exp(\alpha \cdot x^{\beta} + \gamma) \quad (1)$$

where y is the number of items, and x is the rank of items corresponding to Fig. 1. Parameter triplet (α, β, γ) characterizes the modern English language used in the Internet.

We note that this list includes stopwords of the English language, which radically influence the curve at the top ranks. In most computational stylistics papers, stopwords are retained to reveal the style, but in our case, we investigated only the significant words describing the specific theme. By removing stopwords, short words ($|w| \geq 1$), and long words ($|w| \leq 15$) from the English texts, which are later called unnecessary words, the relative frequency of the remaining words changes depending on the context of the text (see *GenerateParams* in subsection B). The term $|w|$ represents the length of words in the number of characters.

A. Characteristics of the Processed Texts

We downloaded 37 ancient texts from Project Gutenberg (see Table I). These texts were split into two main classes (poets and prosaists) and an additional six subclasses devised by Callimachus according to the literary genres of the texts, e.g. tragedians, comic playwrights, speakers, etc. So that the essential content of the texts could be examined, we had to remove unnecessary words and punctuation marks from the original texts.

Word intensity in the merged text (T37) of the analyzed 37 texts is represented in Fig. 2. The fitting equation of the intensity curve was found to conform to the formula below:

$$y(x) = a \cdot x^b + c \qquad (2)$$

where y is the number of items, and x is the rank of items. Equations (1) and (2) are very different. The former equation is exponential to a power function and the latter is a power function. The parameter triplet (a, b, c) characterizes the special features of the ancient English language.

TABLE I.
CLASSES OF ANALYZED ANCIENT TEXTS

	Callimachus' Class	No. of Texts
1.	Poets - Tragedians	8
2.	Poets - Comic playwrights	8
3.	Poets - Epic poets	11
4.	Poets - Lyrical poets	2
5.	Prosaists - Philosophers	5
6.	Prosaists - Speakers	3

Looking at Fig. 1 and 2, we can conclude that modern English tends to use more intensively general words. Those insignificant terms are called stopwords in machine learning-based analysis. Rijsbergen calls attention to Luhn's oeuvre in automatic text analysis, who supposed that word frequencies can be applied to extract words and sentences to represent the content of a document. Let f be the word frequency in a particular position of text and r the rank order of the words (i.e. the order of their frequency of occurrence), then a plot f versus r produces a hyperbolic curve. In addition, this curve presents Zipf's law, which says that the product of word frequencies and their rank order is approximately constant. Luhn applied this law as a null hypothesis to determine the upper and lower cut-offs of the rank order of words. He meant by the significance of the words, their ability to discriminate the topic or content of the text. Using these arbitrarily specified cut-offs, he excluded insignificant words, such as common and rare words, from the rank order of items, and thus he could find those significant words that describe the content of the text [17].

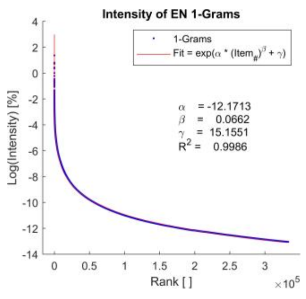


Fig. 1. Frequency of the English words in the Internet texts including all words (logarithmic scale) (Source of raw data: [16])

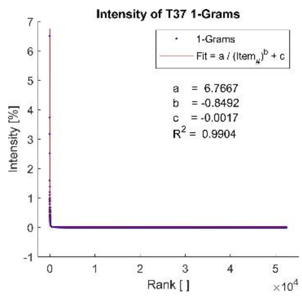


Fig. 2. Frequency of the English words in the analyzed 37 ancient texts including all words (linear scale)

High-level programming languages offer advantageous procedures to eliminate such insignificant elements after converting all the text strings into a list of tokens. Tokens in this context are words or punctuation marks, as well. Examples of such cleaned texts are represented in Fig. 3 (Text1: original author: Aeschylus, translated by Murray, Gilbert, title: Agamemnon, Callimachus' class: Poets – Tragedians; Text4: original author: Apollonius Rhodius, translated by Seaton, R.C., title: The Argonautica, Callimachus' class: Poets - Epic poets). Because of spatial limits in the graphs, just a few elements with the highest ranks are listed from the top list of the tokens. The same methodology of graphical visualization is applied for 2-Grams in Fig. 4, too.

Fig. 3. Word cloud of cleaned Text1 (left) and cleaned Text4 (right).

Stopwords, short words, and long words were removed except archaic English words (i.e. thou, hath, thy, etc.). These cleaned texts contain theme-specific words on the top of the word ranks (see Fig. 3). Word cloud is a representation mode of the token frequency ranks. The larger the size of the characters is, the higher the frequency of the tokens in the text. N-Gram is a group of N tokens ($N = 1, 2, \ldots$) located next to each other in the text. 2-Grams are neighboring word pairs in a fixed order of occurrence. Such examples of 2-Grams for the ancient texts mentioned above are represented in Fig. 4.

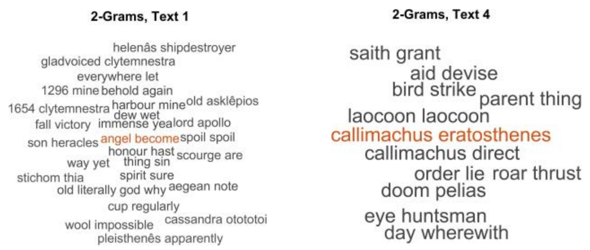


Fig. 4. Bag cloud of 2-Grams of Text1 (left) and Text4 (right).

To execute the removal of the unnecessary tokens from the texts and evaluate artificial intelligence-based procedures, we used embedded packages of the Matlab system: Text Analysis Toolbox and Machine Learning and Deep Learning Toolbox. A list of modern English stopwords we use is integrated into the actual version of the software.

B. Preprocessing of the Texts

To eliminate the effect of stopwords and to generate the rank-fitting parameters for each of $k = 37$ texts, we implemented a cleaning and fitting algorithm given in the following meta code sequence:

```
[a, b, c] = GenerateParams(text1,..., textk):
    global doc, tokenDetails
    for textid = 1:37:
      for partid = 1:10:
        doc = Import(textid, partid)
        doc = Lower(doc)
        doc = TokenizeDoc(doc)
        doc = AddPartOfSpeechDetails(doc)
        tokenDetails = TokenDetails(doc)
        doc = NormalizeWords(doc)
        doc = RemoveStopWords(doc)
        doc = ErasePunctuation(doc)
        doc = RemoveShortWords(doc, max = 2)
        doc = RemoveLongWords(doc, min = 15)
        bag = BagOfWords(doc)
        topBag = TopkWords(bag,noTokens = 1000)
        [a(textid,partid), b(textid,partid), c(textid,partid)] =
        FitCurve(topBag, type = 'hyperbole')
# end GenerateParams()
```

An explanation of tasks executed by each function is enumerated in Table I. The majority of the procedures listed are embedded methods of the Matlab software.

TABLE II.
FUNCTIONS OF TEXT PROCESSING

	Function Name	Effect
1.	*Import()*	Import text part from text
2.	Lower()	Convert string to lower cases
3.	*TokenizeDoc()*	Convert string into tokens
4.	*AddPartOfSpeechDetails()*	Add part of speech details
5.	*NormalizeWords()*	Reduce each token to stem
6.	*RemoveStopWords()*	Delete stopwords of the modern English language
7.	*ErasePunctuation()*	Delete punctuation tokens
8.	*RemoveShortWords()*	Delete short tokens
9.	*RemoveLongWords()*	Delete long tokens
10.	*BagOfWords()*	Generate a bag of words
11.	*TopkWords()*	Generate a top list of words
12.	*FitCurve()*	Fit hyperbole curve to top list

Elements of the preprocessing list were created based on a logical sequence of the tasks. After importing and converting to lower cases, the following steps were executed: i) *Tokenize the text*: This is the process of breaking down a text into individual words or tokens. This step helps in converting the raw text into a structured format where each word is a separate element. It is a fundamental step in text processing. ii) *Add part of speech details*: Knowing the part of speech of each word can provide valuable linguistic information. This information can be useful in later stages of analysis, such as when identifying key terms or relationships between words. iii) *Normalize words*: Stemming involves reducing words to their base or root form. This step helps in consolidating different forms of the same word, which can improve the efficiency of downstream processes by treating variations of a word as a single entity. iv) *Remove a list of stopwords*: Stopwords are common words (e.g., "and," "of," and "the") that often do not carry much meaning and can be considered noise in text analysis. Removing stopwords helps focus on the more meaningful terms in the text. v) *Erase punctuation*: Punctuation marks usually do not contribute much to the semantic meaning of the text and can introduce noise. Removing punctuation simplifies the text and ensures the focus remains on the words. vi) *Remove short words* (with 2 or fewer characters): Very short words, often with two or fewer characters, might not carry significant meaning and can be removed to reduce noise in the data. vii) *Remove long words* (with 15 or more characters): Extremely long words might be outliers or noise in the data. Removing such long words can improve the efficiency of subsequent analysis. viii) *Generate a bag of words*: A bag-of-words representation converts the preprocessed text into a numerical format, representing the frequency of each term in the document. It is a common representation used in many NLP tasks. ix) *Generate a top list of words*: Creating a top list of words helps in identifying the most frequent and potentially important terms in the text. This list can be used for further analysis or visualization. x) *Curve fitting* generates feature triplets (a, b, c) as a compressed property of the text (or subtext).

Note that the function AddPartOfSpeechDetails() retokenizes the text for part-of-speech tagging. This is the reason why it is executed before the NormalizeWords() and stopwords filtering tasks. Stopwords extraction takes place after lemmatization to reduce better the remaining number of words in the dictionary forms. These steps allow tokens to remain in cohesion in the created bag of words (see the order of steps 4, …, 9 in Table II). In addition, Fig. 5 provides further details about the size of the original and cleaned texts. Note that the effect of cleaning tasks results in decreasing text length by one order of magnitude.

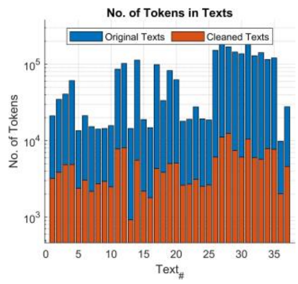


Fig. 5. Reduction of the text length after the preprocessing.

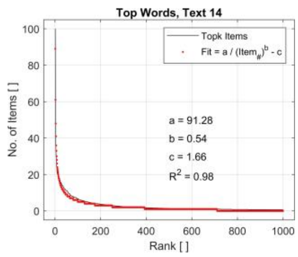


Fig. 6. Number of words in decreasing order of occurrence in Text14.

The resulting two-dimensional arrays a, b, c contain fitting parameters of the text part identified by *textid(1:37)* and *partid(1:10)* indexes, respectively. The function *FitCurve()* executes regression of the rank of the top 1000 words to the hyperbolic equation below:

$$y(x) = \frac{a}{x^b} - c \tag{3}$$

where y is the number of items, and x is the rank of items. Formula (2) and (3) are similar, suggesting no modification of the equation when consistent parts of the text are analyzed. Each parameter triplet *(a, b, c)* characterizes the relative number of words in the corresponding text quantitatively. Variables *doc* and *tokenDetails* have a global scope in the main program and are used for other processing.

An example of the fitting is illustrated in Fig. 6. This text belongs to the Prosaists – Philosophers class. Note that because we have $(b, c) \neq (1,0)$, hence the hyperbole is not a symmetrical curve. We found that all the fitting parameters (b, c) of the 37 texts conforming to the relation (3) are different from $(0, 1)$, respectively. The log-log scale plot of the token intensities vs. ranks is given in Fig. 7. Note that these curves intersect each other in rare cases, so each curve suggests a significant feature of the corresponding text.

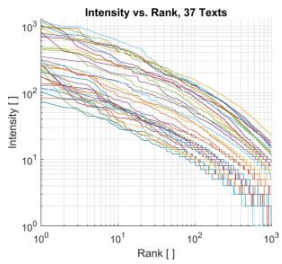


Fig. 7. Log-log plot of the number of words in decreasing order of occurrence in each text.

Representation of the 37 ancient texts in the space of fitting parameters (SFP) can be seen in Fig. 8.

$$SFP = \{(a, b, c) |\ a, b, c, : text\ fit\ parameters\} \quad (4)$$

We realized that texts belonging to the same Callimachus' class are located in compact regions of the SFP.

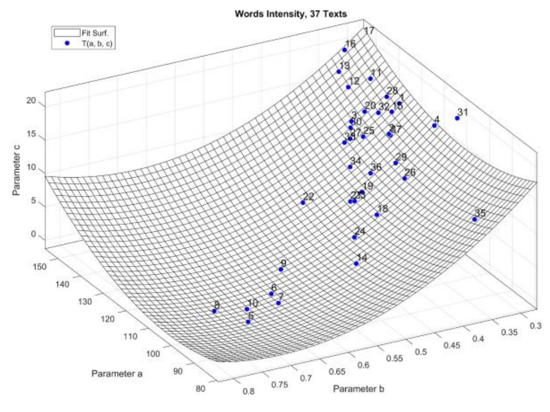


Fig. 8. Location of the texts in the 3D space of parameters (a, b, c).

Some regions on the surface have a small range; others have a more extensive range of parameters. We mention that the region of texts 5-9 belongs to the class 'Poets - Comic playwrights', and that of texts 27-35 belongs to the class 'Poets - Epic poets'. Not all the texts are located homogeneously in this parametric space SFP (i.e. Text35 in the right part of the surface).

We formulate a hypothesis that by dividing each text into ten subtexts, their placement in the parametric space SFP is not changing radically. To validate the hypothesis, we generated $m = 10$ subtexts from each text, and we executed a cleaning and fitting algorithm on each element of the set of $37 \cdot m = 370$ subtexts. The size of each subtext is the tenth part of the parent text, providing 37 different sizes for 370 subtexts.

Fig. 9 proves this statement, but a few subtexts became far positioned in the space SFP (i.e. the top two subtexts). A possible reason for this divergence is that the tenth subtext of some texts contains footnotes generated by the translator, producing an exception of the subtext cohesion.

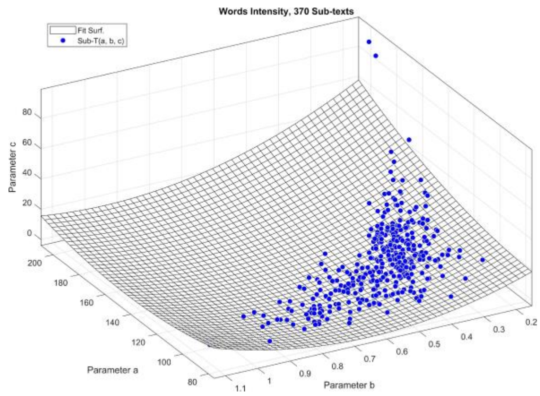


Fig. 9. Location of the subtexts in the space ST of parameters (a, b, c).

We found that the 37 texts and even 370 subtexts are located on a quadrature 'Poly22' type surface given by the following formula:

$$c = p_{00} + p_{10} \cdot a + p_{01} \cdot b + p_{20} \cdot a^2 + p_{11} \cdot a \cdot b + p_{02} \cdot b^2 \quad (5)$$

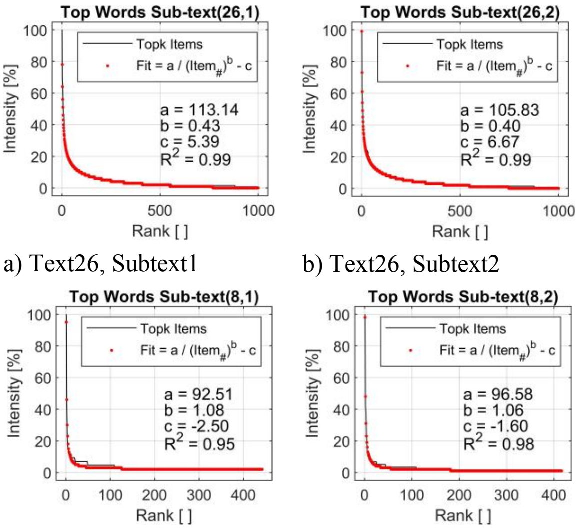
The values of the parameters $p_{ij}, i, j \in \{0, 1, 2\}$ and the coefficient of determination (R^2) of the fitting are given in Table III.

TABLE III.
PARAMETER VALUES OF SURFACE FITTING

Case	p_{00}	p_{10}	p_{01}	p_{20}	p_{11}	p_{02}	R^2
37 Texts	84.14	-0.75	-132.80	0.003	0.11	86.63	0.98
370 Subtexts	36.91	-0.22	-72.30	0.002	-0.30	71.65	0.97

Division of each text into ten subtexts has the following effect: the ranges of variable a remain the same for 37 texts and 370 subtexts: $a \in (80,150)$. In the case of 370 subtexts, the range for variables b and c increases: from $(0.3, 0.8)$ to $(0.3, 0.95)$ and from $(0, 20)$ to $(0, 80)$, respectively.

Examples of hyperbole fitting with a very good coefficient of determination are represented in Fig. 10. We mention that $R^2 > 0.95$ in each case of 370 subtexts. Subtexts 1 and 2 of texts 26 and 8 belong to different Callimachus' classes: Text26: original author: Homer, translated by Butler, Samuel, title: The Iliad of Homer, Callimachus' class: Poets - Epic poets; Text8: original author: Aristophanes, translated by an unknown person, title: The Frogs, Callimachus' class: Poets - Comic playwrights. We note that the rank range of items in subtexts is just 450 because of the reduced size of subtexts. This number is 1000 items for the texts. Another finding is the strong similarity of the subtexts' frequency curves belonging to the same text (subtext(26,1) to subtext(26,2) and subtext(8,1) to subtext(8,2)).



a) Text26, Subtext1 b) Text26, Subtext2

c) Text8, Subtext1 d) Text8, Subtext2

Fig. 10. Top words of Text26 and Text8.

Due to the function *AddPartOfSpeechDetails()* (see Table II) we generated *tokenDetails* objects of each cleaned subtext. This object made it possible for us to have each token category (see Table IV). We generated a relative number of each token category for every subtext producing their histogram.

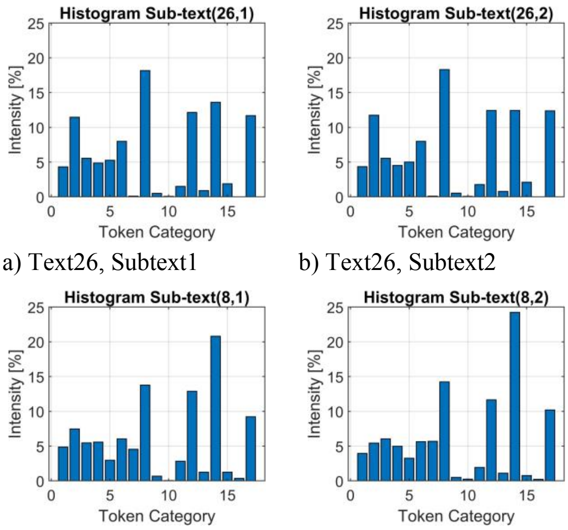
TABLE IV.
TOKEN CATEGORIES AS FEATURES OF THE SUBTEXTS

ID	Token Category	ID	Token Category
1	adjective	9	numeral
2	adposition	10	particle
3	adverb	11	pronoun
4	auxiliary-verb	12	proper-noun
5	coord-conjunction	13	punctuation
6	determiner	14	subord-conjunction
7	interjection	15	symbol
8	noun	16	verb

Histograms of the token categories for four different subtexts are given in Fig. 11. These values are normalized, giving the sum 100%. The value of each token category was found to be in the range $(0, 25\%)$ independently of the subtexts class.

The seventeenth category 'other' is not considered because it is a linear combination of the sixteen categories. We mention that some categories have very low intensity (e.g. symbol, numeral, and particle), and others have high values (e.g. noun and subord-conjunction). Others (e.g. adposition and interjection) have significant fluctuation, providing a difference between the features of the subtexts.

Features are represented by the relative histograms of the token categories. The correlation matrix plots in Fig. 12 and Fig. 13 are composed of cells of 10x10 and 100x100 pixels, respectively, forming squares and long lines (rows and columns) with homogenous colors. White cells on the main diagonal have a value close to 1 and prove the strong correlation between the subtexts of the same text.



a) Text26, Subtext1 b) Text26, Subtext2

c) Text8, Subtext1 d) Text8, Subtext2

Fig. 11. Histogram of token categories of Text26 and Text8.

Dark rectangles or line regions belong to those pairs of texts that are put into different Callimachus' classes. It can be observed in Fig. 12 (in the case of $m = 10$) that the range of the correlation matrix is [0.6, 1], and the mean of the elements is 0.9, proving the strong cohesion of the subtexts in general. The lower correlation is due to the reduced number of tokens of the compared subtexts. In the case of $m = 100$, the range has a lower minimum value, proving that shorter subtexts will not resemble each other. An extreme case is when subtexts contain just one token, and different tokens should belong to different classes, implying a low correlation.

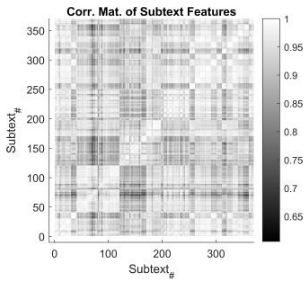


Fig. 12. Correlation matrix of subtext features (m = 10).

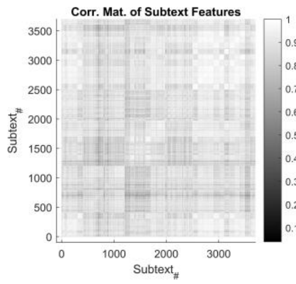


Fig. 13. Correlation matrix of subtext features (m = 100).

C. Neural Network-Based Processing of the Subtexts

Because of the relatively small number of analyzed texts, we divided them into subtexts. We used a significant property of these documents, namely the frequency distribution of words to characterize the texts. We have found that the subtexts belonging to a given text show similar quantitative behavior according to this property. In such an approach, we have $37x10 = 370$ subtexts, each characterized by a feature vector with 16 dimensions. These subtexts are grouped into six

Callimachus' classes according to literary genres. 200, 85, and 85 different subtexts were used for supervised learning with the aim of learning, validation, and testing, respectively. The architecture of the Recurrent Neural Network (RNN) is shown in Fig. 14.

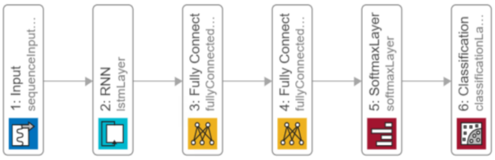


Fig. 14. Architecture of the applied RNN.

The RNN type was Long-Short Term Memory (LSTM), and the applied parameters corresponded to Table V. The LSTM layer (see layer 2 in Fig. 14) was used because RNNs, in general, are more stable in the classification decision than classical NNs without feedback in the structure.

TABLE V.
PARAMETERS OF THE RNN NEURAL NETWORK

Parameter	Value	Parameter	Value
Solver	ADAM	MaxEpochs	30
Gradient Decay Factor	0.90	Mini Batch Size	200
Squared Gradient Decay Factor	0.99	Hidden Units# on L2	100
Initial Learn Rate	0.02	Hidden Units# on L3	100
Gradient Threshold	1	Hidden Units# on L4	6

In the general case of analysis with NNs, statistical features, like accuracy and loss, are evaluated. These features give additional properties of the data analyzed. In the paper we focus primarily on them. The training process of the RNN with the loss and accuracy values in time are represented in Fig. 15. Left and right vertical axes show the loss and the accuracy, respectively. The exponential trend of the loss is proved by a linear trend of the left-hand side axis having a logarithmic scale.

Fig. 15. Loss and accuracy of the training.

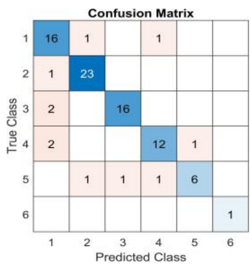


Fig. 16. Loss and accuracy of the training.

Note that the learning process of the RNN was relatively short, only 18.4 sec. We used the remaining 85 subtexts as test inputs of the RNN. It was established that the accuracy of the subtext classification is 87.06%, and the loss is 0.42%. Fig. 16 shows the confusion matrix of the 85 tested subtexts. This result seems very important because each fifth subtext can be identified by the corresponding Callimachus' class without a deeper interpretation of the subtext's meaning.

In Fig. 17 the dependence of the training time and test accuracy is represented vs. the number of subtexts, m ∈{5, 30}. The higher the number of subtexts is, the higher the learning time is. This duration remains under 20 seconds, proving the good performance of the evaluation.

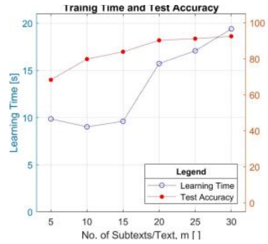


Fig. 17. Training time and test accuracy vs. number of subtexts.

The testing accuracy asymptotically increases with the number of subtexts, if $m \in \{5, 30\}$. Increasing the value of $m > 30$ destroys the testing accuracy. This effect is caused by the very low number of characteristic words (tokens) in the subtexts, removing the subtext from the global context of the original text.

D. Impact of the Communication Errors on the Subtexts Classification

Classical computer networks transfer text content in error-free mode. New transfer mechanisms offer quality-based services with nonzero but low error rates because of the significantly reduced price of these later transfers. We simulated the effect of transfer quality by introducing noise on the LSTM RNN performance during the classification of the subtexts. The noise was generated by replacing the original characters of the words in random positions with the next character in the alphabet. The last character, 'z' was replaced cyclically by the first character, 'a'. In most character replacements, the new words have no real meaning, so these words are assigned to the token category 'Other'. We should mention a few situations when the new word gets true meaning (i.e. 'war' has the third character 'r' replaced with the next letter 's', creating the word 'was'). It will probably influence the selection of the token category, too (see Table IV). The likelihood of these cases is very slight. Therefore, we did not consider them explicitly. According to Table II, this pollution of the original texts was executed before the text processing. The effect of this noisy modification was evaluated in the function of character error rate (CER), p, and the number of subtexts, m. For a relatively high character error rate ($p = 16\%$), the correlation between subtexts remains over 78% for a split ratio of the texts in $m = 20$ subtexts (see Fig. 18).

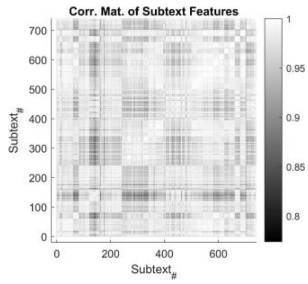
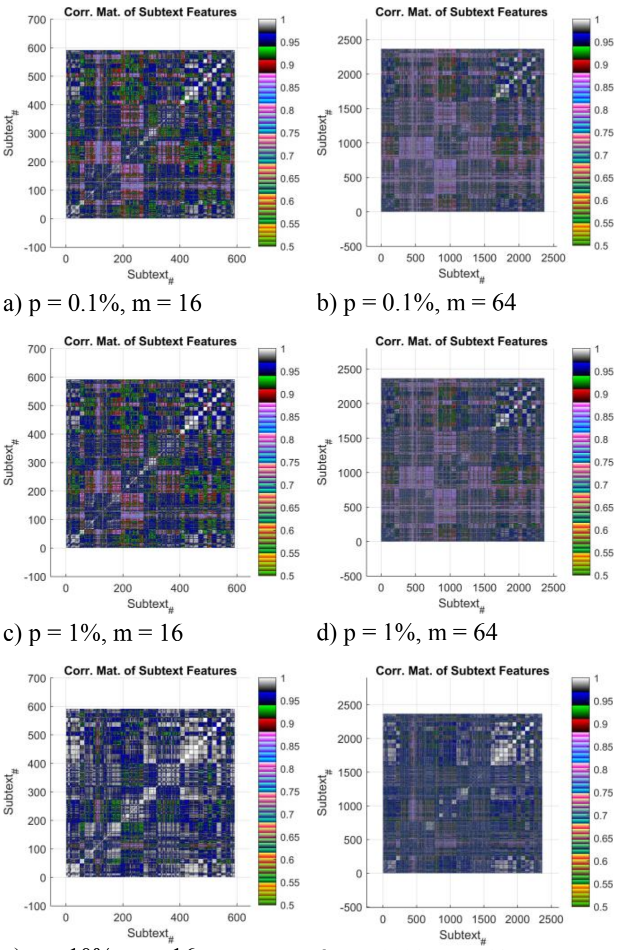


Fig. 18. Correlation matrix of subtext features (p = 16%, m = 20).

Based on Fig. 12, 13, and 18, we found that the minimum value (0.78%) of the correlation ranges is given by case $m = 20$. It is caused by the reduction of tokens in the subtexts corresponding to the error rate of the characters with probability, $p = 16\%$. In this sense, the effect of the pollution is reduced.

Fig. 19 illustrates the correlation matrixes between subtexts of different split ratios and error rate cases. The left (a, c, e) and right (b, d, f) hand side figures belong to the split ratio $m = 16$, and $m = 64$, respectively. Figure rows show character error rates of 0.1%, 1%, and 10%. The patterns of the two columns are very different because the number of subtexts is distinct. This implies considerable differences regarding the number of collected tokens from the subtexts, causing incomparable patterns of the corresponding correlation matrixes.

We realized that for low error rates, $p \in \{0.1\%, 1\%\}$ the patterns for $m = 16$ are very similar but not identical. The same property exists in the cases of $m = 64$. When the character error rate is large, $p = 10\%$, the pattern of the correlation matrix is significantly distinguishable from the cases of the same split ratio (see pairs c-e and d-f of Fig. 19, respectively).



a) p = 0.1%, m = 16 b) p = 0.1%, m = 64

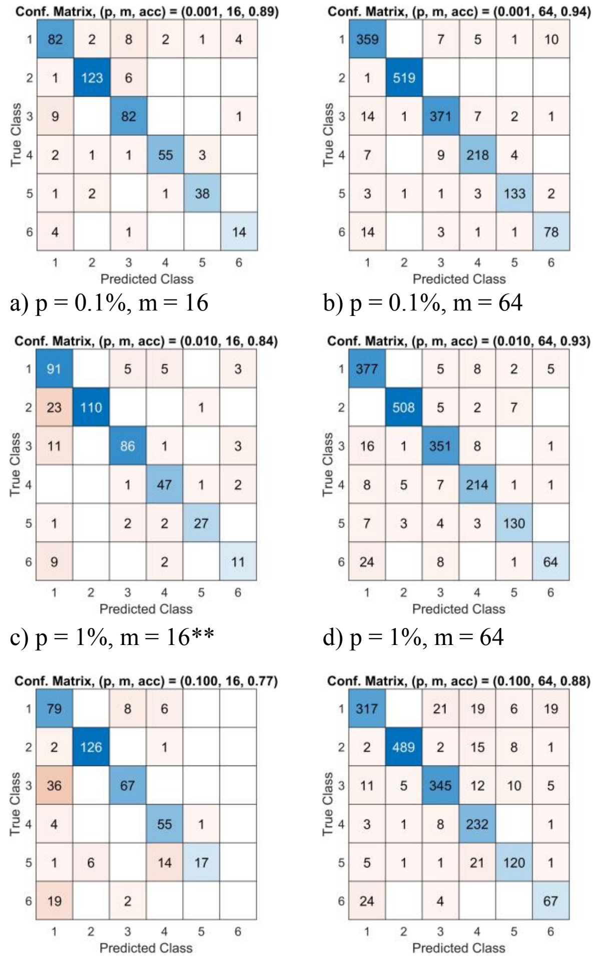
c) p = 1%, m = 16 d) p = 1%, m = 64

e) p = 10%, m = 16 f) p = 10%, m = 64

Fig. 19. Dependence of the correlation matrixes on the character error rate and no. of subtexts.

The white color of the pixels of the figures reflects a high correlation, being close to value 1. Pattern e of Fig. 19 stands for the same number of subtexts as patterns a and c. However, because it has a larger character error rate, the size of the subtexts is smaller, resulting in a smaller number of representative tokens enrolled into categories of Table IV. These remaining tokens correlate more to the common context, implying white-colored clusters with high cohesion of the neighboring subtexts.

The number of six Callimachus' classes remains unchanged, but the character error rate influences the accuracy of the subtext classification made with the LSTM neural network. The higher the value of the parameter p is, the lower the accuracy of the RNN classification of the subtexts is (see confusion matrixes a-c-e and b-d-f of Fig. 20, respectively). The higher the number of subtexts is, the higher the accuracy of the RNN classification becomes. This rule cannot remain valid for any parameter m value, because of the limited number of tokens in the subtexts. Other situations may occur when we have only one token in each subtext. The majority of the tokens alone have no context (i.e., 'war', 'god', 'king', etc.); just a few of them may be considered to have slight context (i.e., 'Zeus', 'Agamemnon', 'Prometheus'). This effect needs further study.



a) p = 0.1%, m = 16 b) p = 0.1%, m = 64

c) p = 1%, m = 16** d) p = 1%, m = 64

e) p = 10%, m = 16 f) p = 10%, m = 64

Fig. 20. Dependence of the confusion matrixes on the character error rate and no. of subtexts.

As the character error rate rises beyond 10%, the accuracy of the subtext classification based on LSTM RNN starts to decrease. Fig. 21 shows the dependence of the classification accuracy of the neural network on the character error rate in the case when $m = 50$.

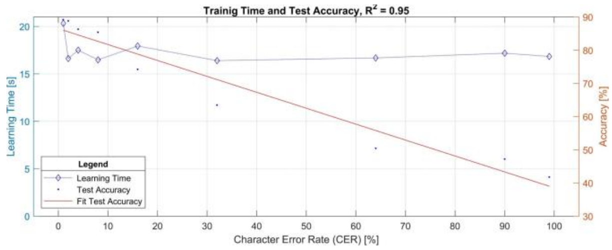


Fig. 21. Training time and test accuracy vs. character error rate (m = 50).

We determined that the accuracy of the ancient Greek text classification executed with the LSTM neural network when the texts are split into $m = 50$ subtexts is approximately linear dependent on the character error rate. The explicit equation is the following:

$$Accuracy\ [\%] \cong 0.9 - 0.5 \cdot p \qquad (6)$$

This equation has a minimum value of 40% when the character error rate approaches value 1. It means that if we strongly pollute these texts, the classification accuracy remains considerable. This property is typical of the ancient Greek texts. It remains an open question to address what is the case with the other textual data depending on their subject.

IV. Conclusions

In this paper we proved that the ancient Greek texts translated into archaic English language have an essential feature: extracting statistical properties of the most frequent 1000 grammatical tokens from the subtexts of 37 texts makes it possible to identify Callimachus' class of subtexts with the probability of 87.06%. The identification of the subtexts' class was accomplished by a Long-Short Term Memory type recurrent neural network. This result is worthy of note for two reasons: the meaning of the verbal content is not required to be considered in the classification process of texts; the learning time of the recurrent neural network takes less than one minute on an ordinary desktop computer. Because of the relatively small number of ancient texts accessible in English, the question of the usage of this proposed methodology can be addressed in detail in another research paper, and the methodology can be refined, if necessary. A possible continuation of this research work is the determination of the plot dynamicity of the texts and the classification of the texts based on their dynamicity. We plan to represent this new dimension of the texts in a multimodal way (e.g., image and sound). Long-Short Term Memory type of neural networks can classify subtexts of ancient Greek literary works even in the case of missing words. This method can be applied to help interpret the old manuscripts affected by the physical deterioration of their material. A possible continuation of the research work may be finding the appropriate methods and features to differentiate between lyric and prose form translations of the same text.

Acknowledgment

This work has been supported by the QoS-HPC-IoT Laboratory and project TKP2021-NKTA of the University of Debrecen, Hungary. Project no. TKP2021-NKTA-34 has been implemented with the support provided from the National Research, Development and Innovation Fund of Hungary, financed under the TKP2021-NKTA funding scheme.

References

[1] H. F. Hsieh, and S. E. Shannon, "Three approaches to qualitative content analysis," *Qualitative health research* vol. 15, no. 9, pp. 1277–1288, 2005.

[2] K. Krippendorff, Content analysis: An introduction to its methodology. California: Sage, 2018.

[3] K. A. Neuendorf, The content analysis guidebook. 2nd ed. London: Sage, 2016

[4] J. Kasperiuniene, M. Briediene, and V. Zydziunaite, "Automatic content analysis of social media short texts: Scoping review of methods and tools," *Computer Supported Qualitative Research. WCQR 2019. Advances in Intelligent Systems and Computing*, vol. 1068, A. P. Costa, L. P. Reis, A. Moreira, Eds. Cham, Switzerland: Springer, cop. 2020, pp. 89–101.

[5] J. Singh, G. Singh, R. Singh, P. Singh, "Optimizing accuracy of sentiment analysis using deep learning based classification technique," *Data Science and Analytics. REDSET 2017. Communications in Computer and Information Science*, vol. 799, B. Panda, S. Sharma, N. Roy, Eds. Singapure: Springer, 2018, pp. 516–532.

[6] J. R. Macnamara, "Media content analysis: Its uses, benefits and best practice methodology," *Asia Pacific Public Relations Journal*, vol. 6, no. 1, pp. 1–34, 2005.

[7] Y. H. Chang, C. Y. Chang, and Y. H. Tseng, "Trends of science education research: An automatic content analysis," *Journal of Science Education and Technology* vol. 19(4), pp. 315–331, 2010.

[8] P. Baranyi and Á. Csapó, "Definition and synergies of Cognitive Infocommunications," *Acta Polytechnica Hungarica*, vol. 9, no. 1, pp. 67–83, 2012.

[9] P. Baranyi, A. Csapó, and Gy. Sallai, Cognitive Infocommunications (CogInfoCom). Springer International, 2015.

[10] Carl Vogel and Anna Esposito, "Interaction Analysis and Cognitive Infocommunications," *Infocommunications Journal*, vol. XII, no 1, March 2020, pp. 2–9. DOI: 10.36244/ICJ.2020.1.1

[11] I. Boda, M. Bényei, and E. Tóth, "New dimensions of an ancient Library: the Library of Alexandria," in *CogInfoCom 2013. Proc. of the 4th IEEE International Conference on Cognitive Infocommunications*, (Budapest, Hungary, December 2-5, 2013,) pp. 537–542.

[12] *MaxWhere VR Even more*. [Online]. (2023-04-28) Available: http://www.maxwhere.com/ .

[13] I. Boda and E. Tóth, "Text-based approach to second language learning in the virtual space focusing on Callimachus' life and works," in *P. Baranyi Ed. CogInfoCom 2019. Proc. of the 10th IEEE International Conference on Cognitive Infocommunications*, (Naples, Italy, October 23-25, 2019,) pp. 439–444.

[14] M. A. Korteby, Z. Gál, and P. Polgár, "Multi dimensional analysis of sensor communication processes," *Annales Mathematicae et Informaticae*, vol. 53, pp. 169–182, 2021.

[15] A. Aldabbas, Z. Gál, K. M. Ghori, M. Imran, M. Shoaib, "Deep Learning-Based Approach for Detecting Trajectory Modifications of Cassini- Huygens Spacecraft," *IEEE ACCESS*, vol. 9, pp. 39 111–39 125, 2021.

[16] *Kaggle Database: 1/3 Million Most Frequent English Words on the Web*. [Online]. (2023-04-28) Available: https://www.kaggle.com/rtatman/english-word-frequency

[17] C. J. van Rijsbergen, Information Retrieval. 2nd ed. London: Butterworths, 1979. [Online]. (2023-04-28) Available: http://www.dcs.gla.ac.uk/Keith/Preface.html

[18] Van Halteren, H., “Source language markers in EUROPARL translations,” in *Proc. of the 22nd International Conference on Computational Linguistics (Coling 2008)*, 2008, pp. 937–944.
[19] Lynch, G. and Vogel, C., “The translator’s visibility: Detecting translatorial fingerprints in contemporaneous parallel translations,” *Computer Speech & Language*, vol. 52, pp. 79–104., 2018.
[20] Koppel, M. and Ordan, N., “Translationese and its dialects,” in *Proc. of the 49th annual meeting of the association for computational linguistics: Human language technologies*, 2011, pp. 1318–1326.

Zoltan Gal is currently a full professor at the Faculty of Informatics, University of Debrecen, Hungary. He earned MSc in electrical engineering and computer science from the Technical University of Timisoara, Romania and PhD in informatics sciences from the University of Debrecen. His scientific interest is focused on distributed processing and communication systems, sensor technologies and services in the Internet of Things.

He was the CIO of his institute for 20 years and developed the university-level metropolitan area highs-speed data network and services with over 10k Internet nodes. He is Cisco Certified Network Professional lecturer since 1999 and taught over five hundred network professionals in the field. Starting in 2015 he is head of the Centre of High-Performance Computing at his university. He is an IEEE member since 1996 and published over one hundred twenty scientific conferences and journal papers: He supervises his own R&D&I project called QoS- HPC-IoT Laboratory.

Erzsébet Tóth is an assistant professor at the Faculty of Informatics, University of Debrecen, Hungary. She had a degree in English language and literature and library and information science in 1995. In 2008 she obtained her Ph.D degree in information science and technology. In the doctoral dissertation she investigated the evaluation of search engine performance. She is involved in a virtual library project that focuses on the presentation of digitized library content in three-dimensional space, and she studies the enhanced possibilities of the English language teaching and learning in virtual learning environment. Since 2022 she has been an IEEE member.

Finding the contextual impacts on Students' Mathematical performance using a Machine Learning-based Approach

Zakaria Khoudi*, Mourad Nachaoui, and Soufiane Lyaqini

***Abstract*—An extensive dataset for examining Moroccan eighth-grade pupils' mathematical prowess was made available by the 2019 Trends in Mathematics and Science Study (TIMSS). The TIMSS 2019 public dataset contained 8390 Moroccan students, who were the subject of this research. Based on how well they could solve mathematical problems, the participants were split into 3108 high achievers and 5282 poor achievers in the mathematics phase of the exam. This study aimed to pinpoint the essential environmental elements affecting eighth-grade pupils' mathematical abilities. In order to do this, the research used cutting-edge machine learning methods, particularly the efficient distributed gradient boosting toolkit XGBoost. From a vast collection of 700 possible components, this strategy proved critical in identifying the most relevant variables. These factors included a broad spectrum of components at the student, teacher, and school levels. After a thorough investigation, 12 critical contextual factors distinguishing between arithmetic prodigies and average performers were successfully found. The discovery of these critical characteristics has significant implications for future instructional efforts, especially in improving high school pupils' mathematical proficiency. Knowledge of these factors may assist educators and policymakers in creating focused interventions and pedagogical approaches that enhance mathematics performance and comprehension. This research emphasizes how complex mathematics accomplishment is and how crucial it is to approach educational planning holistically. Identifying and addressing these critical environmental elements can significantly enhance students' mathematics achievements at a crucial juncture in their academic development.**

***Index Terms*—Contextual factors, Machine learning, Mathematics performance, Moroccan students, TIMSS 2019**

I. Introduction

THE International Association for the Evaluation of Educational Achievement (IEA) created the international research project known as TIMSS, which is conducted every four years. For further information, check [4]. TIMSS aims to track mathematical and scientific success trends and investigate the effects of various educational factors, curricula, and resources on student accomplishment. The most recent worldwide comprehensive study of math performance was TIMSS 2019. Data from 250.000 students, 30.000 teachers, and 8.000 administrators were used to evaluate grade 8 mathematics in the 39 participating nations, including Morocco. International achievement exams, particularly math and science, have frequently shown the low performance of Moroccan students. Lower test scores are thought to reflect lower levels of educational quality and human capital [18], which have a significant and negative impact on worker outcomes in the labor market as well as the nation's economic growth and human development [15], [20], [24], [1]. As a result, Morocco was ranked lower in the rankings of participating countries in the most recent International Mathematics and Science Study TIMSS 2019. Researchers may now examine the many variables that affect student accomplishment thanks to the widely accessible, comprehensive, and policy-relevant indicators evaluated by TIMSS 2019. This research can aid in the decision-making of policymakers, teachers, and students.

This research investigates the potential causes of Moroccan eighth-graders poor mathematical performance in TIMSS 2019 by using machine learning (ML) methods. Because there is currently a shortage of research on applying machine learning (ML) techniques for international assessment studies, we look into characteristics of the teacher, school, and student variables that have a high significance in predicting the mathematics test results of Moroccan college students. Researchers may compare the outcomes in terms of student accomplishment thanks to TIMSS, which also offers information on the impact of policies and practices in the educational systems of each participating country [25]. Standard statistical techniques, including regression analysis, multilevel modeling, and component analysis, are frequently used when examining student performance in science and mathematics. [19], [28], [31], [34], [37]. However, as previously indicated, these techniques have inherent shortcomings, mainly when dealing with strangely distributed data. Additionally, it might be challenging to make accurate forecasts [30] when the situation is highly complicated. Machine learning techniques are effective at finding outliers [10] and may be used to get around some of the drawbacks of these conventional approaches. This project aims to incorporate all TIMSS-provided student, teacher, and school variables in 3 models to identify significant predictors of students' mathematical success using machine learning methods such as SVM, XGBoost, and Random forests.

The remaining sections are arranged as follows: section II displays related works to this research, section III provides an overview of the data, and an explanation of ML approaches,

Zakaria Khoudi and Mourad Nachaoui are with the Equipe Mathématiques et Interactions, Faculté des Sciences et Techniques, Sultan Moulay Slimane University, Beni Mellal, Morocco.
(E-mail: zakaria.khoudi@usms.ma; m.nachaoui@usms.ma)
Soufiane Lyaqini is with the LAMSAD Laboratory, Ecole Nationale des Sciences Appliquees, Hassan First University, Settat, Morocco.
(E-mail: lyaqini.soufiane@gmail.com)
* Whom correspondence should be addressed

DOI: 10.36244/ICJ.2024.5.2

section IV shows and analyzes estimate findings, and section VI wraps up the article.

II. Related Works

Most studies described above were created to collect data on assessing education systems. Researchers may compare/improve educational policies and their outcomes in terms of success by using educational data, which includes assessments of various nations' educational performance using tools like TIMSS. Modeling the TIMSS data using machine learning techniques was the subject of much significant research. Finding the variables affecting students' mathematical achievement was the goal of Hammouri [14]. A student's success in mathematics is determined by their attitude, accomplishment, confidence in their mathematical abilities, and sense of the importance of mathematics, according to the findings produced using the TIMSS dataset. Liu and Meng [21] investigated the TIMSS 2003 dataset and used these variables to compare the mathematical awareness of high and poor achievers in East Asia and America. For the TIMSS 2011 data, Askin and Gokalp [2] looked at the variables that affect students' performance in their academic endeavors. They employed LR, and ANN approaches to evaluate the performance of prediction and classification. The confidence of the students was shown to be the most important component. Topçu, Erbilgin utilized data from the TIMSS 2011, and Arkan [34] to look at what influences Turkish and Korean students' achievement in science and mathematics. The educational ramifications of their results were also highlighted. Data from TIMSS 2011 were subjected to DT, NB, LR, and ANN applications by Klç-Depren, Askin, and Oz [11]. They sought to identify the top algorithm for categorizing Turkish eighth-graders based on various performance indicators for their proficiency in mathematics. Filiz and Oz [12] used the TIMSS 2015 scientific data to apply the EDM approach. They discovered the elements that contribute most to scientific accomplishment. Baranyi and Gilanyi [3], Chmielewska [8], [9] unveil the notion of "Mathability". This idea highlights how artificial intelligence and human cognitive capacities may be integrated, mainly when comprehending and solving mathematical problems. They demonstrate how improvements in information and communications technologies may improve human mathematic abilities. The authors highlight the possibility of integrating cognitive processes and technology tools in improving mathematical abilities by demonstrating math ability in complicated problem-solving, such as utilizing Maple to solve mathematical equations and inequalities. This study is in line with comprehending the variables that affect students' performance in mathematics, especially as it relates to how technological interventions and cognitive capacities can work together to enhance learning outcomes in mathematics. This is an essential area of research for studies that use machine learning techniques to predict and analyze academic performance. Contribute most to the advancement of science.

Since little research in Morocco focuses on contextual factors impacting children's mathematics competency, this study fills the knowledge gap by utilizing machine learning methods to examine the impact of the student, teacher, and school variables on high and poor-performing children. Two research concerns are addressed in this study:

- Can context help Moroccan eighth-graders with high mathematics levels and those with low mathematics levels be distinguished? If so, what are they?
- What contextual elements should the ideal feature set have that would influence Moroccan pupils' mathematics competence, both high and low?

III. Material and Methods

A. Data description

The study's dataset was obtained from TIMSS 2019 database conducted by the IEA (International Association for the Evaluation of Educational Achievement) https://timss2019.org/international-database/, and it used Moroccan student's grade 8 data files after combining the student, teacher, and school data using IDB Analyzer (version 5.0) offered by IEA, initially 8458 entries almost equal in terms of the number of males and females (Figure 1) with 700 variables (Figure 2) were made.

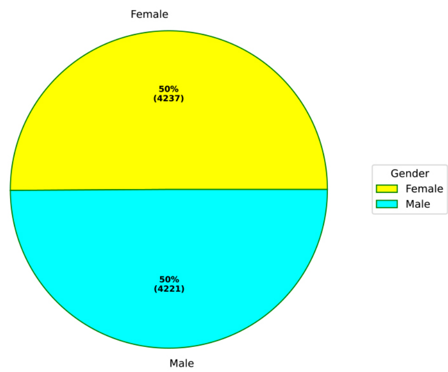


Fig. 1. Gender in the dataset

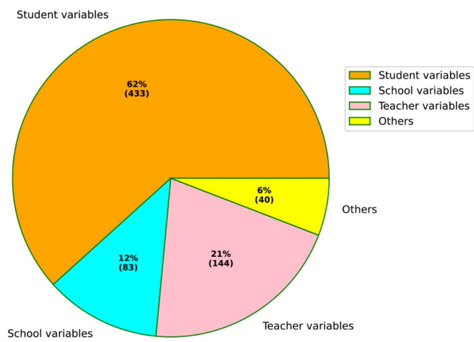


Fig. 2. Structure of dataset variables

Special Issue
of the Infocommunication Journal

These variables are divided into four parts below:

- The variables beginning with BC****** refer to a school background (e.g., BCBG16I, BCBG19, BCBG15D, Etc.), a total of 83 variables.
- The variables beginning with BS****** refer to a student's background (e.g., BSBG10, BSBG11B, BSBG13C, BSBM19B, Etc.), a total of 433 variables.
- The variables beginning with BT****** refer to a teacher background (e.g., BTBM17CA, BTBM17CC, BTBG12F, Etc.), a total of 144 variables.
- IDs, the number of teachers, weights and file maintenance are the remaining variables, totalling 40.

1) Students background: As preliminary observations on student background variables:

- We have randomly chosen a sample of 2000 students. Also, eight variables have been chosen. There is a clear association between teaching clarity across all scientific subjects (possible link to instructional quality at the school level). (Figure 3)
- A wide range in the number of books in the house, more than 50 per cent of students have a few books at home. (Figure 4)
- The most significant level of parental education varies greatly, although low level "Some Primary, Lower Secondary or No School" is the most typical. (Figure 5)

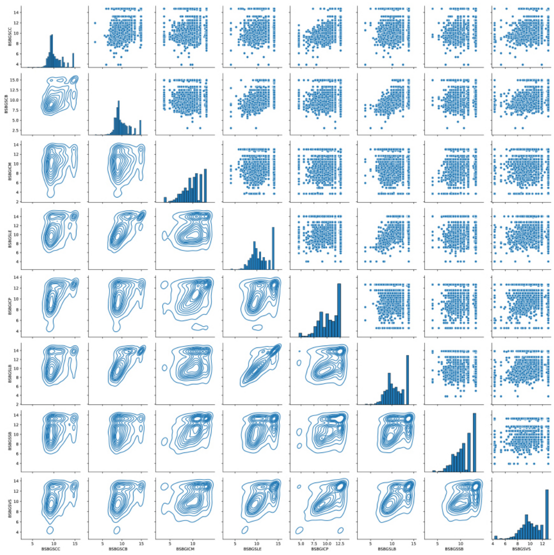


Fig. 3. Correlations between a few of the demographic scores for students' scores

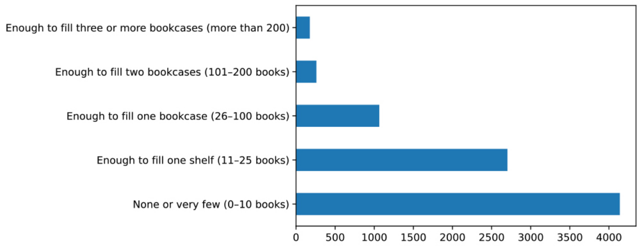


Fig. 4. Each student's home library's number of books

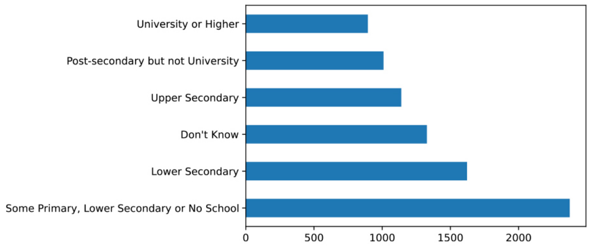


Fig. 5. The highest degree of education attained by the parents of the student

2) Teachers background: The dataset contains 260 math teachers divided into 251 high schools; therefore, each student is taught by one teacher. As an initial finding on the background factors of math teachers are

- Roughly Normal age distribution among teachers, the majority of teachers are young. Their ages range from 25 to 39. (Figure 6)
- Bachelor's degrees are where the majority of education for teachers is found. (Figure 7)
- There are noticeably more male than female teachers. Males outnumber females by more than 50 per cent. (Figure 8)

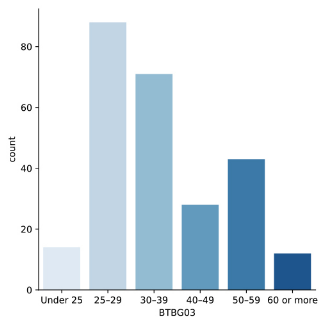


Fig. 6. The age range of math instructors

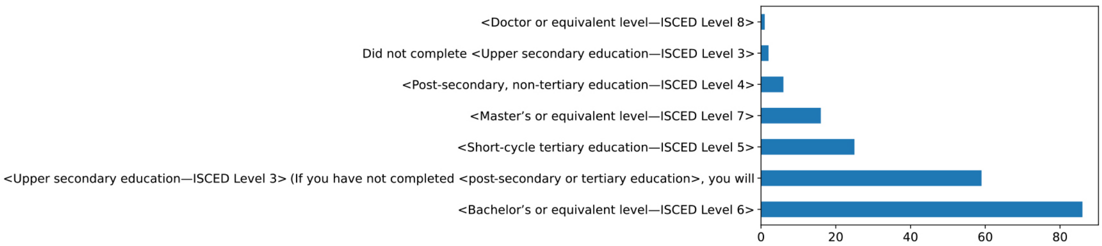


Fig. 7. The frequency of math instructors' academic successs

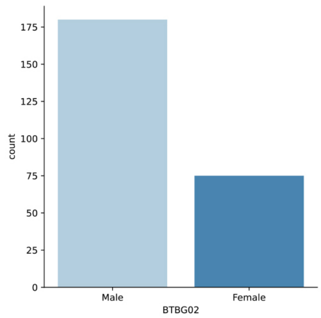


Fig. 8. Math instructors' gender distribution

3) Schools background: The initial dataset includes 251 private and public Moroccan high schools, randomly selected from rural and urban areas. Figure 9 highlights the relationship between discipline, academic competencies, and shortage of equipment in these high schools. As preliminary notices, we found:

- There does not seem to be a relationship between school disciplinary issues and an emphasis on academic performance or a lack of resources.
- The lack of resources for math and science is closely connected because schools seem to have difficulties with both subjects. Meanwhile, some other schools do not have such difficulties.
- But few of these spectrums have little problems with discipline and resources.

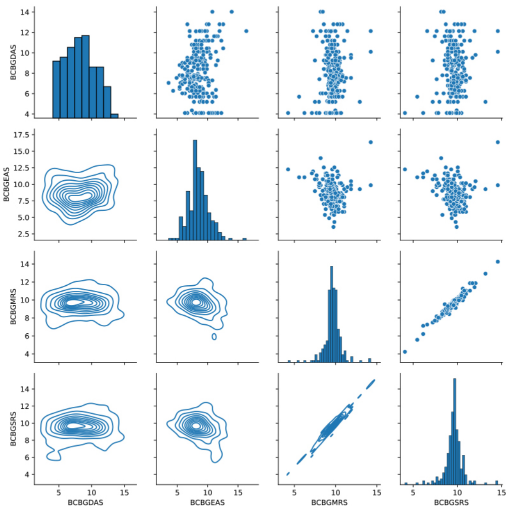


Fig. 9. Shortages, academics, and disciplinary relationships

B. Response variable

TIMSS awards 5 PVs (plausible values) to a student's academic achievement. PVs are then classified into five levels: 1 (Below Low), 2 (Low), 3 (Intermediate), 4 (High), and 5 (Extremely High) (Advanced), and these classed variables are referred to as categorical benchmark variables using the five categories of benchmark variables and a majority vote, a single class was created for each student's math achievement. For example, if a student's benchmark variables (BSMIBM01 through BSMIBM05) were 3, 1, 1, 1, 3, the student's class was coded as 1. After combining the 8458 Moroccan 8th graders using the IDB Analyzer (version 5.0), 68 of them had ties (Figure 10). After removing the ties, the final sample for this study was 8390 students.

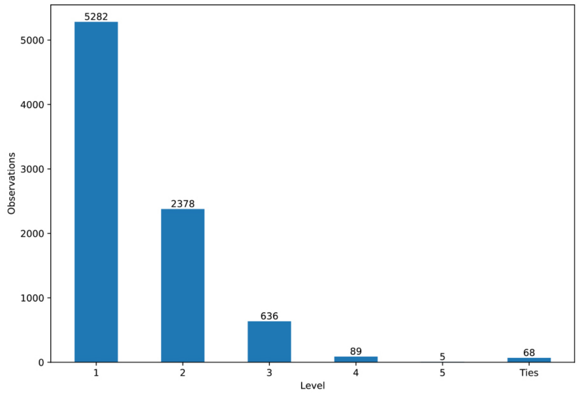


Fig. 10. TIMSS 2019 Moroccan 8th Grade Math Majority Vote Results

Because the proportions of the levels were severely imbalanced, the final four levels were compacted in this study. The study's response variable was whether the student attained the 'Low' level (Level 1: coded as 0) or not (Levels 2, 3, 4, 5, coded as 1). There were 5282 and 3108 students in each group (Figure 11).

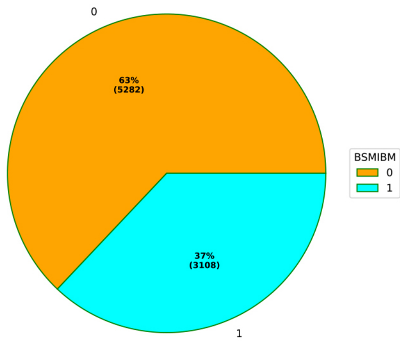


Fig. 11. Response variable categories

C. Data Pre-processing

The combined dataset originally had 700 variables, and 300 of them were eliminated, as shown below:

- 41 variables relevant to IDs (e.g., IDPOP, IDSTUD, IDCLASS, Etc.), weights (e.g., total school weight, total class weight Etc.), file maintenance (e.g., VERSION, IDB Identifier, Etc.), and redundant variables (e.g., ITSEX) were removed from the explanatory variable pool.
- 61 variables with 100% missingness (e.g., BSBS22D, BSBS22E, BSDGSEC, BSDGSCS, etc.) were eliminated.
- 29 numerical scale score variables of category indices for constructs such as 'Students Like Learning Mathematics Lessons' (e.g., BSBGSLM, BSBGICM, BSBGDML, Etc.) were deleted because numerical scale scores and categorical indices transmit essentially the same information.

- 95 benchmark variables and PVs were removed, except for the newly generated benchmark variable, BSMIBM, which functioned as the study's response variable. Including these academic performances, factors would have dominated the model, providing little meaningful information for predicting students' math proficiency.
- Missing values include omitted or invalid, logically inapplicable, and not administered. The "omitted" replies resulted from respondents' carelessness or refusal to answer the question. Because the educational systems of the TIMSS participating nations were so diverse that the replies were mainly "not applicable-administered". Each variable's missing rate was determined, and 74 variables with missing rates of more than 10% were deleted from the dataset.
  The cleaned dataset had 400 variables from 8390 students, including 205 students, 122 teachers, and 73 school variables (Figure 12). Only 11 were continuous variables with hours and numbers as replies (e.g., BTBG11, BTBM14, BCBG18, BTDMNUM, whereas the rest 389 were Likert-type scaled.

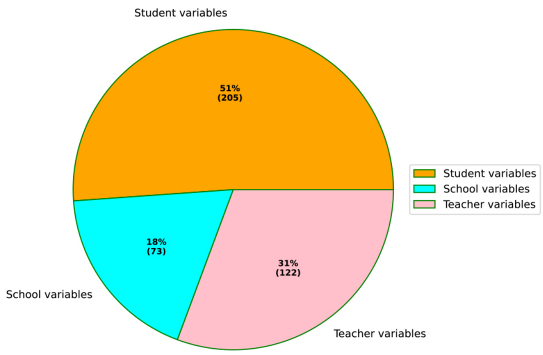


Fig. 12. The final structure of dataset variables

For missing data imputation, we used the SimpleImputer function from sklearn.impute package, which successfully retained all 8390 observations with 400 variables. The SimpleImputer begins by replacing missing data with mode values. The 8390 observations are almost equal in terms of the number of males and females (Figure 13).

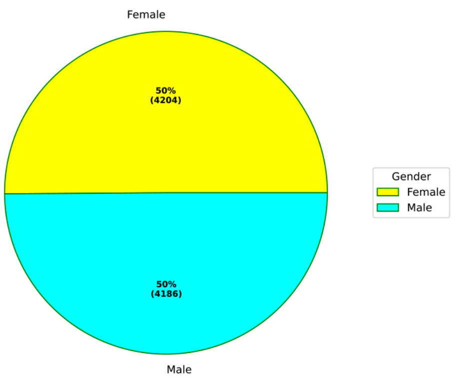


Fig. 13. Gender in the dataset after cleaning

D. Classification Models

After pre-processing and preparing the real datasets, we obtained knowledge of the data set. Important variables affecting students' mathematical achievement may be predicted using a variety of classifiers. Our experiment uses the supervised models Support Vector Machine (SVM), Random Forest, and XGBoost since no machine learning approach consistently produces the best predictions.

SVM is a method for creating classifiers [16], [22],[27]. Its objective is to create a judgment boundary between two groups that enables labels to be predicted from one or more feature vectors. The hyperplane, a judgment boundary, is angled to be as close as feasible to one of the classes' closest data points. The closest-together points form help vectors.

Random forest [5] is a Machine Learning method for handling classification and regression issues. It is based on ensemble learning, a method that combines several classifiers to provide answers to complicated problems. Many decision trees are used in it. With bagging or bootstrap aggregation, the algorithm's created forest is trained. We selected it due its frequent use for feature selection in a data science workflow [26], [13], [17]. The average impurity reduction derived from every decision tree in the forest may determine how important a feature is, whatever the data's linear or non-linear nature. It is true (linearly inseparable).

XGBoost [7] is an application of gradient-boosted decision trees designed to quickly and accurately resolve various data science challenges. When dealing with the bias-variance trade-off, boosting algorithms are pretty helpful. Boosting handles both the elements of bias and variance, in contrast to bagging algorithms that solely correct for excessive variance in a model. One of the top machine learning models at the moment. Due of its speed, effectiveness, and scalability [36], [6], [7]. Extracting feature significance from the XGBoost model is simple compared to other machine learning models since it is effectively an ensemble of decision trees. Because of this, we decided to use XGBoost in this research to determine feature relevance.

The observations were divided into random training and test data sets, with 7:3 being the standard ratio. The test data was utilized for generalizing the model once it had been evaluated using the training data.

In particular, the response variable BSMIBM was used as a stratifying variable to maintain the ratio of "Low" to "Others" in the training and test datasets. The student counts for the training and test data sets are shown in (Table I) for each level.

TABLE I
TRAINING AND TEST DATA

	0(Low : 63%)	1(Others : 37%)
Data(n = 8390)	5286	3104
Training data (n = 5873)	3700	2173
Test data (n = 2517)	1586	931

The ability to distinguish between instances and controls is a feature of XGboost. The optimal values for the hyperparameters of the best model are found using grid search with 10-fold Cross Validation [29] within the train split. The recommended hyperparameter settings for XGBoost are learning_rate = 0.3, max_depth = 10, and n_estimators = 400. Then, we take the critical characteristics out of XGboost. The primary function of XGBoost, a component of its Python library, is used to assess the applicability of each feature. By adding a feature, the method determines the average training loss decrease for each splitting.

IV. Experimental results

The results reported in this section are based on the three classification models previously covered in the "Classification Models" section.

A. Testing and evaluation

We employed a variety of assessment metrics, or indicators, to assess and comprehend the performance of the models. A confusion matrix [32] is a simple technique to see how well a model works. For our purposes, the positive class represents high-level students, while the negative class represents low-level students.

TABLE II
Confusion matrix in general

Actual Positive	True Positive (TP)	False Negative (FN) Type I error
Actual Negative	False Positive (FP) Type II error	True Negative (TN)
	Predicted Positive	Predicted Negative

We assess the effectiveness of our prediction models using four ML model assessment indicators [35]. Accuracy and F1-score are the two measures used to evaluate classifier performance. The percentage of all correct predictions serves as a measure of accuracy. The method for calculating accuracy is shown in this equation:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

True Positive is abbreviated as TP. True negative is TN, whereas false positive is FP and FN for "false negative". The following equation represents the F1-score. For a classification task, it reflects the harmonic means of the values for accuracy (Predicted features) and recall (sensitivity):

$$F1 = \frac{2 * (Precision * Recall)}{Precision + Recall}$$

Area Under Curve (AUC) [23]: The typical rule of thumb is using a probability threshold of 0.5 for classification predictions in binary classification issues. Using a different threshold might be preferable in a few circumstances when this one would not hold. The most popular way to display a binary classifier's performance at various thresholds is through a Receiver Operating Characteristic (ROC) curve. The True Positive Rate is plotted against the False Positive Rate to achieve it. The false-positive rate is computed as (1 - Specificity). The ROC plot enables us to calculate the Area Under the Curve by estimating the probability that a classifier will score a randomly picked positive instance higher than a randomly selected negative one (commonly abbreviated as simply the AUC).

B. Results

The results of our performance assessment have been presented in Table III, Figures (14,15,16,17,18,19) so that we can identify the key indicators of students' mathematical competency. Our investigation employed three different models: SVM, Random Forests, and XGBoost. However, the scores for these two models' accuracy might have been higher, indicating they may have had tremendous success fixing the present problem. XGBoost models, on the other hand, produced excellent results, with an accuracy of **81.92** and an F1- Score of **74.42**. These findings suggest that XGBoost models might be better suited to identify the ideal proportion of crucial contextual variables that affect eighth-grade kids' mathematical aptitude.

TABLE III
Results of performance assessment

Models	Accuracy	F1-Score
SVM	65.87	16.03
Random Forest	79.18	65.39
XGBoost	**81.92**	**74.42**

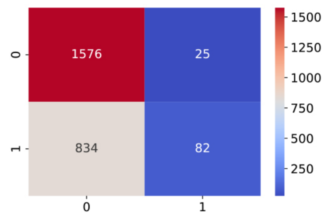


Fig. 14. SVM confusion matrix

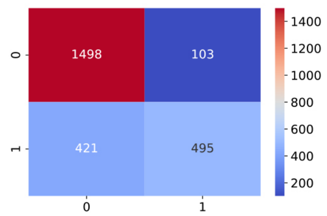


Fig. 15. Random Forests confusion matrix

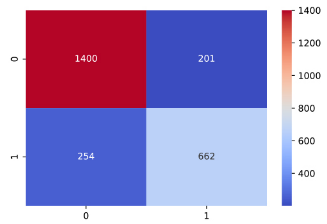


Fig. 16. XGboost confusion matrix

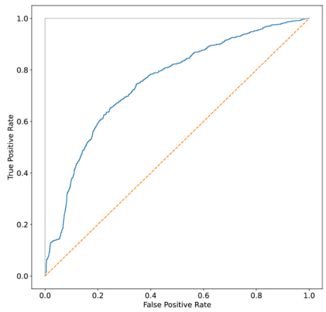


Fig. 17. SVM AUC-ROC Curve

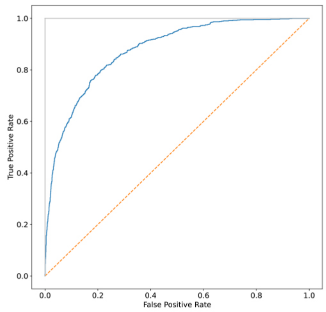


Fig. 18. Random Forests confusion matrix

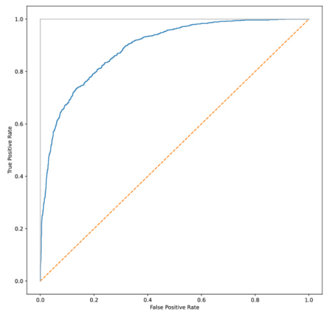


Fig. 19. XGboost confusion matrix

C. Feature selection

A machine learning approach known as feature selection involves selecting a smaller sample of critical features from a more extensive range of unimportant or irrelevant data while minimizing information loss. It is often used when dealing with many characteristics and few instances. Feature selection may accelerate processing and increase prediction accuracy by decreasing the quantity of data that the algorithm must evaluate [33].

In machine learning, "feature importance" is used to pinpoint the features or variables in a dataset that significantly affect the target variable. This method is essential for understanding the interactions and behavior of variables inside a model and aids in identifying the variables that are most important for result prediction.

In our work, feature selection assigns a score to each input feature in a particular model using measurements of feature relevance as criteria. Higher scores imply features that significantly impact the prediction model, representing each feature's relative relevance. There are other ways to calculate feature significance, but XGBoost yields the best results. The significance of several characteristics is shown in Figure 20 from XGBoost, which shows how well the model predicts feature importance. The top 12 traits are highlighted in the graphic, which lists them in decreasing order of importance.

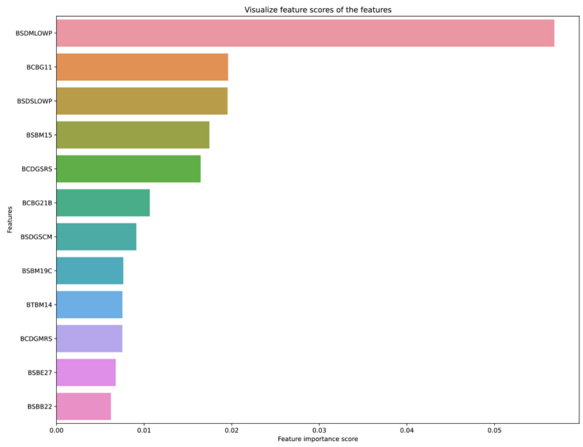


Fig. 20. XGBoost Features Importance

V. Discussion

The experiment's findings demonstrate that XGBoost offers the essential components required for the analysis. The differences between BSDMLOWP, BCBG11, BSDSLOWP, BSBM15, BCDGSRS, BCBG21B, BSDGSCM, BSBM19C, BTBM14, BCDGMRS, BSBE27, and BSBB22 are in the sequence in which they are provided. When determining the mathematical competency of Moroccan students, these 12 qualities provide the best selections for predictor variables since they regularly appear in XGBoost and have high relevance. Table IV provides descriptions of these salient traits. These essential characteristics may be used as predictor variables, according to the descriptions that have been given.

TABLE IV
Description of essential characteristics

Feature	Description	Type
BSDMLOWP	MATHEMATICS ACHIEVEMENT TOO LOW FOR ESTIMATION	1: Yes; 2: No
BCBG11	CLASSROOM LIBRARIES	1: Yes; 2: No
BSDSLOWP	SCIENCE ACHIEVEMENT TOO LOW FOR ESTIMATION	1: Yes; 2: No
BSBM15	WORK ON YOUR OWN	1: Every or almost every lesson; 2: About half the lessons; 3: Some lessons; 4: Never
BCDGSRS	INSTRUCTION AFFECTED BY SCIENCE RESOURCE SHORTAGE	1: Not Affected; 2: Affected; 3: Affected A Lo
BCBG21B	QUALIFICATIONS IN EDUCATIONAL LEADERSHIP/ ISCED 7	1: Yes; 2: No
BSDGSCM	STUDENT CONFIDENT IN MATHEMATICS	1: Very Confident in Mathematics; 2: Somewhat Confident in Mathematics; 3: Not Confident in Mathematics
BSBM19C	MATHEMATICS NOT MY STRENGTH	1: Agree a lot; 2: Agree a little; 3: Disagree a little; 4: Disagree a lot
BTBM14	TIME SPENT MATH INSTRUCTION MINUTES	1 to 1800
BCDGMRS	INSTRUCTION AFFECTED BY MATHEMATICS RESOURCE SHORTAGE	1: Not Affected; 2: Affected; 3: Affected A Lot
BSBE27	HOW OFTEN CONDUCT EXPERIMENTS IN EARTH SCIENCE	1: At least once a week; 2: Once or twice a month; 3: A few times a year; 4: Never
BSBB22	HOW OFTEN CONDUCT EXPERIMENTS IN BIOLOGY	1: At least once a week; 2: Once or twice a month; 3: A few times a year; 4: Never

To go over the outcomes in light of the two study-related research concerns:

- **Q1: Can context help Moroccan eighth-graders with high mathematics levels and those with low mathematics levels be distinguished? If so, what are they?**

Context may assist in separating Moroccan eighth-graders who do well in mathematics from those who do not. According to our research, the following contextual factors are crucial for drawing this distinction:

1) Student factors: It was discovered that seven different student-related factors significantly affected mathematics ability. These include kids' accomplishment levels in science and math, their self-perceptions of their mathematical prowess, the regularity with which they work independently in math class, and their confidence in mathematics, among other things. The study showed that variables about students were more critical in influencing academic achievement than those about instructors and educational institutions. This highlights how important it is to tailor instructional tactics to each learner's unique requirements and features.
2) School and Teacher Variables: The impact of teacher and school characteristics was also assessed in this research. These include classroom libraries, educational leadership credentials, and the effects of scientific and math resource shortages.
3) Teacher Variables: An assessment of the amount of time spent on math teaching was part of the research, and the results showed how important it is to students' academic achievement.
4) Overall Effect and Educational Implications: According to the study, variables that affect students significantly influence their academic achievement more than those that affect educators and educational institutions. Given the present focus on teacher and school accountability in education, this emphasizes the need for more research to examine and identify attributes of teachers and schools.

the following contextual factors help differentiate Moroccan eighth-graders with high and low mathematics achievement levels: confidence in mathematics, frequency of independent math work, perception of math as a strength or weakness, time spent on math instruction, the impact of resource scarcity, and overall mathematics achievement level. These elements provide a thorough grasp of the pupils' mathematical aptitude.

- **Q2: What contextual elements should the ideal feature set have that would influence Moroccan pupils' mathematics competence, both high and low?**

Various contextual factors that include many facets of the student's educational environment, personal attitudes, and accessible resources should be included in the optimal feature set for impacting Moroccan students' high and poor mathematical competency. The following components are essential, according to the TIMSS 2019 study:

1) Student Confidence in Mathematics (BSDGSCM): This variable assesses students' confidence in their aptitude for mathematics. Students' motivation, enthusiasm, and achievement in mathematics may all be significantly impacted by developing their confidence. Students with differing degrees of mathematical proficiency might benefit from tailored interventions and teaching strategies.
2) Frequency of Independent Math Work (BSBM15): Students' quantity of independent math work in the classroom is significant. Students' critical thinking and quantitative skills may be improved by promoting self-directed learning and problem-solving.
3) Viewing Mathematics as a Weakness (BSBM19C): Students' performance may be impacted by how they see themselves in mathematics, specifically if they consider math a personal weakness. Addressing this perspective and making it a strength may be helpful for kids with limited mathematical proficiency.

4) Time Spent on Math Teaching (BTBM14): The amount of time dedicated to math teaching is crucial. Ensuring enough time for teaching may enhance students' mathematical understanding and performance. Educators and schools must take this into account when developing a curriculum.
5) The Effect of Mathematical Resource Scarcity (BCDGMRS): The quality and quantity of mathematical resources greatly influence how well students are taught mathematics. Improving students' comprehension and interest in mathematics requires addressing the need for more resources, particularly for those with lower competency levels.
6) Mathematics Achievement Level (BSDMLOWP): This variable aids in identifying pupils whose competency in the subject may be too low to evaluate reliably. Giving these pupils focused assistance may help them become more proficient in mathematics.

Including these contextual factors in instructional techniques and interventions may produce a more customized and successful strategy for raising Moroccan students' mathematical proficiency. It entails concentrating on each student's attitudes, confidence, learning environment, and content and teaching.

The study's findings have implications for educational policy and practice as they may help educators and policymakers create focused interventions and resources that will enhance student outcomes, especially in mathematics. Having a better understanding of the main performance-influencing variables may aid in the development of instructional and support systems that are more successful.

Several vital elements emerge in the concept of "Mathability" as defined by Baranyi and Gilanyi [3], along with the field of cognitive infocommunications. Firstly, cognitive infocommunications seeks to understand the coevolution of artificial and natural cognitive processes, shedding light on how students learn and understand mathematical concepts and how artificial systems like machine learning algorithms can be utilized to assess and predict student performance. Secondly, Baranyi and Gilanyi's notion of "Mathability" refers to the capability to simulate and enhance human mathematical skills, aligning closely with the study's focus on factors affecting students' math performance. Insights into learning environments, achievement levels, and student confidence can inform the development of artificial systems or educational aids replicating and augmenting these aspects. Thirdly, applying these concepts in educational technology can lead to creating tools that support personalized learning, addressing individual student needs and boosting overall math skills. In summary, viewing the study's findings on factors impacting the mathematical performance of Moroccan eighth-graders through the lenses of mathability and cognitive infocommunications can pave the way for more effective teaching methodologies and technological aids that nurture and support mathematical competence.

This research examined 700 student, teacher, and school features from the TIMSS 2019 to predict the mathematical performance of Moroccan eighth-graders. It identified the top 12 variables using XGBoost. Only a small number of teacher and school features were also found by XGBoost, even though all seven variables were related to students, indicating factors affecting students had a more significant impact on their academic performance than those affecting teachers and schools. This was expected, but more research is needed to study and pinpoint teacher and school features, especially in the current climate of increased focus on teacher and school responsibility. Education researchers must provide teachers and schools with research-based assistance since they are now more than ever held accountable for their student's academic success. That was this empirical study's primary goal.

VI. Conclusion

This research thoroughly examines important factors affecting Moroccan students' mathematical skills, including student, teacher, and school characteristics. It highlights twelve key predictor characteristics, such as a poor estimate of math performance, individual math work, and insufficient science and math teaching resources. These variables are essential markers as there is a strong correlation between them and pupils' arithmetic performance. Additionally, the research reveals other noteworthy characteristics that provide insightful information to scholars, professionals, and decision-makers who want to improve Moroccan students' mathematical skills by implementing methods and interventions specific to these impactful elements.

References

[1] Noam Angrist, Simeon Djankov, Pinelopi K Goldberg, and Harry A Patrinos. Measuring human capital using global learning data. *Nature*, 592(7854):403–408, 2021.

[2] Oykum Esra Askin and Fulya Gokalp. Comparing the predictive and classification performances of logistic regression and neural networks: a case study on timss 2011. *Procedia-Social and Behavioral Sciences*, 106:667–676, 2013.

[3] Peter Baranyi and Attila Gilányi. Mathability: Emulating and enhancing human mathematical capabilities. In *2013 IEEE 4th international conference on cognitive Infocommunications (CogInfoCom)*, pages 555–558. IEEE, 2013.

[4] Antoine Bodin and IREM de Franche-Comté. Ce qui est vraiment évalué par pisa en mathématiques. ce qui ne l'est pas. un point de vue français. *Bulletin de l'APMEP*, 463:240–265, 2006.

[5] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[6] Jason Brownlee. *XGBoost With python: Gradient boosted trees with XGBoost and scikit-learn*. Machine Learning Mastery, 2016.

[7] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794, 2016.

[8] Katarzyna Chmielewska. From mathability to learnability. In *2019 10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, pages 513–516. IEEE, 2019.

[9] Katarzyna Chmielewska. Towards "learnability". *Acta Polytechnica Hungarica*, 19(1), 2022.

[10] Karel Dejaeger, Frank Goethals, Antonio Giangreco, Lapo Mola, and Bart Baesens. Gaining insight into student satisfaction using comprehensible data mining techniques. *European Journal of Operational Research*, 218(2):548–562, 2012.

[11] Serpil Kılıç Depren, Öyküm Esra Aşkın, and Ersoy Öz. Identifying the classification performances of educational data mining methods: A case study for timss. Educational Sciences: *Theory & Practice*, 17(5), 2017.

[12] Enes Filiz and Ersoy Öz. Finding the best algorithms and effective factors in classification of turkish science student success. *Journal of Baltic Science Education*, 18(2):239–253, 2019.

[13] Baptiste Gregorutti, Bertrand Michel, and Philippe Saint-Pierre. Correlation and variable importance in random forests. *Statistics and Computing*, 27(3):659–678, 2017.
[14] Hind Hammouri. Attitudinal and motivational variables related to mathematics achievement in jordan: Findings from the third international mathematics and science study (timss). *Educational research*, 46(3):241–257, 2004.
[15] Eric A Hanushek and Dennis D Kimko. Schooling, labor-force quality, and the growth of nations. *American economic review*, 90(5):1184–1208, 2000.
[16] Marti A. Hearst, Susan T Dumais, Edgar Osuna, John Platt, and Bernhard Scholkopf. Support vector machines. *IEEE Intelligent Systems and their applications*, 13(4):18–28, 1998.
[17] Mia Huljanah, Zuherman Rustam, Suarsih Utama, and Titin Siswantining. Feature selection using random forest classifier for predicting prostate cancer. In *IOP Conference Series: Materials Science and Engineering*, volume 546, page 052 031. IOP Publishing, 2019.
[18] Zakaria Khoudi, Mourad Nachaoui, and Soufiane Lyaqini. Identifying the contextual factors related to the reading performance of moroccan fourth-grade students from a machine learning-based approach. *Education and Information Technologies*, pages 1–27, 2023.
[19] Serpil Kilic and Öyküm Esra Askin. Parental influence on students' mathematics achievement: The comparative study of turkey and best performer countries in timss 2011. *Procedia-social and behavioral sciences*, 106:2000–2007, 2013.
[20] Jong-Wha Lee and Robert J Barro. Schooling quality in a cross–section of countries. *Economica*, 68(272):465–488, 2001.
[21] Shujie Liu and Lingqi Meng. Re-examining factor structure of the attitudinal items from timss 2003 in cross-cultural study of mathematics self-concept. *Educational Psychology*, 30(6):699–712, 2010.
[22] S Lyaqini, M Nachaoui, and A Hadri. An efficient primal-dual method for solving non-smooth machine learning problem. *Chaos, Solitons & Fractals*, 155:111754, 2022.
[23] Caren Marzban. The roc curve and the area under it as performance measures. *Weather and Forecasting*, 19(6):1106–1114, 2004.
[24] David Mitch. Education matters. Global schooling gains from the 19th to the 21st century. By Robert J. Barro and Jong-Wha Lee. New York: Oxford University Press, 2015. pp. xi, 289, 34.95, hardcover. *The Journal of Economic History*, 76(4):1271–1273, 2016.
[25] Ina VS Mullis, Michael O Martin, Pierre Foy, and Alka Arora. *TIMSS 2019 international results in mathematics*. ERIC, 2020.
[26] Md Mursalin, Yuan Zhang, Yuehui Chen, and Nitesh V Chawla. Automated epileptic seizure detection using improved correlation-based feature selection with random forest classifier. *Neurocomputing*, 241:204–214, 2017.
[27] Mourad Nachaoui, Soufiane Lyaqini, and Marouane Chaouch. Indicating if water is safe for human consumption using an enhanced machine learning approach. *Statistics, Optimization & Information Computing*, 11(1):70–81, 2023.
[28] Oliver Neuschmidt, Juliane Barth, and Dirk Hastedt. Trends in gender differences in mathematics and science (timss 1995–2003). *Studies in Educational Evaluation*, 34(2):56–72, 2008.
[29] Isaac Kofi Nti, Owusu Nyarko-Boateng, and Justice Aning. Performance of machine learning algorithms with different k values in k-fold cross-validation. *J. Inf. Technol. Comput. Sci*, 6:61–71, 2021.
[30] Muhammad A Razi and Kuriakose Athappilly. A comparative predictive analysis of neural networks (nns), nonlinear regression and classification and regression tree (cart) models. *Expert systems with applications*, 29(1):65–74, 2005.
[31] James B Schreiber. Scoring above the international average: A logistic regression model of the timss advanced mathematics exam. 2000.
[32] Thomas R Shultz, Scott E Fahlman, Susan Craw, Periklis Andritsos, Panayiotis Tsaparas, Ricardo Silva, Chris Drummond, PL Lanzi, J Gama, RP Wiegand, et al. Confusion matrix. *Encyclopedia of machine learning*, 61(8):209–209, 2011.
[33] Jiliang Tang, Salem Alelyani, and Huan Liu. Feature selection for classification: A review. *Data classification: Algorithms and applications*, page 37, 2014.
[34] Mustafa Sami Topçu, Evrim Erbilgin, and Serkan Arıkan. Factors predicting turkish and korean students' science and mathematics achievement in timss 2011. 2016.
[35] Ž Vujović et al. Classification model evaluation metrics. *International Journal of Advanced Computer Science and Applications*, 12(6):599–606, 2021.
[36] Yan Wang and Xuelei Sherry Ni. Axg boostrisk model via feature selection and bayesian hyper-parameter optimization. *arXiv preprint ArXiv:*1901.08433, 2019.
[37] Ludger Wößmann. Educational production in east asia: The impact of family background and schooling policies on student performance. *German Economic Review*, 6(3):331–353, 2005.

Zakaria Khoudi has received master's degree in information and communication technology from Abdelmalek Essaadi University. Now, he is a PhD student in computer science at Sultan Moulay Slimane University. His research interests are in data science, machine learning and data mining.

Mourad Nachaoui has received his PhD in 2011 from Nantes University. He is an professor at Sultan Moulay Slimane University. He was a guest professor at the Laboratory of Mathematics Jean Leray, Nantes University, from 2013 to 2019. His main research interests are applied mathematics, inverse problems, data sciences and computer sciences.

Soufiane Lyaqini has received his PhD from Sultan Moulay Slimane University. Now, he is an associate professor at Hassan 1 University. His research interests are in machine learning, data mining and optimization.

Assessing Memory Colors of University Students

Cecilia Sik-Lanyi, Bence Halmosi, Jinat Ara, Judit Szűcs and Tibor Guzsvinecz

***Abstract*—Human perception is complex, and can be influenced by several factors. The study's purpose is to understand these factors, thus, an application was developed. This application examines memory colors in the CIELAB color space. Memory colors can be affected by nationality, virtual reality game playtime, and the presence or absence of an image cue. The memory colors of banana and orange had the highest agreement between the students, while the memory colors of river and grass had the largest dispersion. Virtual reality games influenced the memory color of grass the most. It can also be concluded that without an image cue, different colors were selected in 70.90% of the cases.**

***Index Terms*—cognitive aspects of virtual reality, cognitive infocommunications, color perception, cross-cultural study, human-computer interaction, memory color, virtual reality.**

I. Introduction

THE electromagnetic radiation that reaches our eye and produces sensations is called color stimulus. Consequently, this sensation produces the perception of colors in our brain. Three components of color perception are distinguished in the scientific literature: brightness, hue and colorfulness [1]:

- Brightness: the sensation can be almost blinding strong, medium or dim and dark;
- Hue is usually shown as a hue circle, where four distinguishable different areas are red, yellow, green and blue. A yellow hue can be reddish or greenish, but can never be bluish; a green hue can be yellowish or bluish, but can never be reddish, and so on [2]. It is possible to define the hues between two fundamental (or unique) hue, as e.g. green, bluish green, greenish blue, blue. Sometimes for the hues that are somewhere in the middle between two unique hues special names is used:
  a. Yellow – red: orange
  b. Green – yellow: lemon
  c. Blue – green: cyan or turquoise
  d. Red – blue: magenta or purple
- Colorfulness has been divided by MacDonald into a five-value scale: Starting with gray (achromatic) up to the most brilliant color. If no hue can be determined for a color then it is gray (or white or black). Thus, with increasing colorfulness one can speak about a gray, grayish, moderately vivid, vivid and very vivid color perception.

Therefore, for example, the communication of color perception can be made in the following form: the person's complexion who stands in front of us is medium bright, moderately vivid and reddish yellow [3].

However, as computer scientists, the human visual system can be described as follows: visual perception is the act of observing patterns and objects through vision or visual input. At the highest level, vision systems are almost the same for humans, animals, and most living organisms. They consist of a sensor that captures an image and a brain that processes and interprets an image [4]. Then, it outputs a prediction based on the data extracted from the input image, as shown in Figs. 1 and 2. In other words, the human visual system perceives depth through a sequential process involving the analysis of color, motion, form, and ultimately, the construction of a three-dimensional perception of the visual scene.

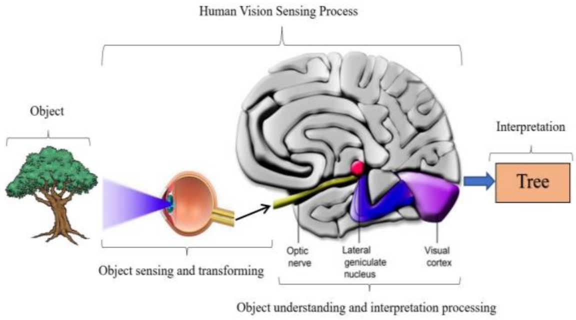


Fig. 1. The human vision system

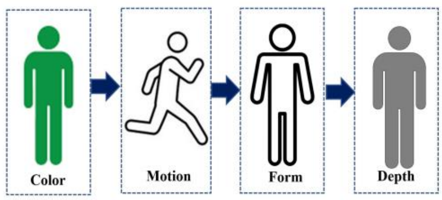


Fig. 2. How depth is formed.

Submitted on 7/1/2023

Cecilia Sik-Lanyi is with the Department of Electrical Engineering and Information Systems, Faculty of Information Technology, University of Pannonia, Veszprem, Hungary (e-mail: lanyi.cecilia@mik.uni-pannon.hu). She is the corresponding author.
Bence Halmosi is with the Department of Electrical Engineering and Information Systems, Faculty of Information Technology, University of Pannonia, Veszprem, Hungary (e-mail: halmosi.bence@mik.uni-pannon.hu).
Jinat Ara is with the Department of Electrical Engineering and Information Systems, Faculty of Information Technology, University of Pannonia, Veszprem, Hungary (e-mail: jinat.ara@mik.uni-pannon.hu).
Judit Szűcs is with the Department of Information Technology and its Applications, Faculty of Information Technology, University of Pannonia, Zalaegerszeg, Hungary (e-mail: szucs.judit@zek.uni-pannon.hu).
Tibor Guzsvinecz is with the Department of Information Technology and its Applications, Faculty of Information Technology, University of Pannonia, Zalaegerszeg, Hungary (e-mail: guzsvinecz.tibor@zek.uni-pannon.hu).

DOI: 10.36244/ICJ.2024.5.3

A memory color is the general color of an object that we recall due to our experience with it. For instance, most individuals are aware that ripe bananas are yellow in color. Our ability to consistently replicate an object's color in our minds increases with frequency of exposure [5]. However, memory color does not always match the original color. According to research, these colors are stored by the brain as being brighter and more saturated than the colors of the actual objects. This is most likely a result of the fact that our brains store them as the most aesthetically pleasing colors of an object. Although there are undoubtedly differences between people in terms of memory colors, the majority of us share a common memory color for everyday objects like the human face or other natural objects [6].

Lightness, chroma, and hue shifts can be observed in short-term memory, and neither sensory mechanisms nor differences in adaptation can account for these shifts [7-10]. These shifts can be explained by the cognitive effect hypothesis as an exaggeration. Focal colors, prototypical colors, and color regions have various effects on color perception. This results in various degrees of color memory shifts [11]. As can be seen, the previously mentioned phenomena and effects are cognitive. If we want to assess these in the digital world, the Cognitive InfoCommunications (CogInfoCom) field of research provides the perfect toolkit for this study. CogInfoCom is an interdisciplinary field that examines the relationship between human, information, and communication technologies [12-15]. Thus, it gives an opportunity to investigate several human factors using modern cognitive IT methods. Among others, human-computer interaction as well as human vision are investigated in the field of CogInfoCom [16-21].

Thus, the purpose of this article is to investigate the influence of various factors, including nationality, virtual reality (VR) game playtime, and the presence or absence of image cues on memory colors within the CIELAB color space. By examining the impact of these factors on the perception of memory colors, the study aims to contribute to a deeper understanding of human perception and its practical implications, particularly in the context of virtual environments and color cognition.

In order to present this pilot study, this article is structured as follows. Section II presents related studies. The materials and methods are shown in section III. The results are detailed in section IV, while they are discussed in section V. Conclusions are drawn in section VI.

II. Related studies

There are numerous studies on the subject of memory colors. Bartleson tested 50 participants to find the memory color for 10 everyday objects using the Munsell color scale, which consists of hue, value, and chroma. The conclusion drawn from researching ellipses is that there is some consistency in color choice. This implies that the hues are not noticeably different. After taking hue and saturation into account, he came to the conclusion that two-thirds of the color selections were 2.44 patches apart at most [6]. In a study from 2006, participants were asked to change the hue of several fruits so that they would appear grayscale. When the fruit's original hue was changed to the opposite color, the subjects thought it was gray. For instance, to make bananas appear gray, subjects gave them a blueish shade (which is the opposite of yellow) [22]. These findings show that visual memory has a significant impact on how color is formed in our brain that it is not entirely based on the perceived information [23].

Studies have also revealed that the colors used in virtual reality video games frequently differ from those that may be observed in the real world, for instance, the grass in certain games is significantly darker [24]. Another investigation into memory colors was done in 2009. Sik-Lanyi determined that virtual reality video game addicts associated darker green color with grass [25]. Nationality-related research was conducted by Tarczali, although only the memory colors of Korean and Hungarian testers were compared [26]. According to the study's findings, "the effect of a visual cue on memory plays a significant role in the study of color shifts, and its presence or absence can alter memory in different ways throughout observation" [27]. A study on the impact of culture on long-term memory span was carried out in 2014 in seven countries (Belgium, Brazil, Colombia, China, Hungary, Iran, and Taiwan). On a calibrated monitor, individuals in each country were shown every-day objects in more than 100 different colors. The testers were tasked to judge whether an object resembled the original object when they saw it in a particular color. According to statistical analysis, the averages of countries and the averages of all exams were significantly different [28].

A test software was developed to test the memory colors associated with brands [29]. In this software, participants had to color logos of brands. The results show that if participants had to remember more colors, they had selected more wrong colors when coloring logos of brands.

III. Materials and methods

The CIELAB color space was used since it is closer to human vision [30]. Color can be described in the CIELAB color system with the following triplets: $L*$ is called lightness, and has values between 0 and 100, while $a*$ and $b*$ are color scales. Both range between the values of -128 and 127. Red-green colors can be found on the $a*$ scale, while yellow-blue can be observed on $b*$. This color system is shown as a three-dimensional body, with the lightness (the brightness of surface colors), chroma (representing colorfulness) axes and hue circle (or $a*$, $b*$ axes to describe chroma and hue) where the hue is measured as a hue angle.

Participation in the tests was voluntary, and 54 people took part in them. Participants included both Hungarians and foreigners. A room that was fully dark and free of any outside light was used for the testing. The only source of illumination was the desktop display. A 60 cm distance separated each participant from the monitor. The tasks were explained before the tests began. The participants then had to fill out a form with their age, gender, nationality, and how many hours they spend each week playing virtual reality games. According to the responses, students were divided into four groups: university students from Hungary, international students, those who play video games frequently (i.e. more than 20 hours a week), and those who do not. The number of participants grouped by nationality was the following: 25 Hungarians, 8 Chinese, 7 Russians, 5 Spanish, 3 Jordanians, 2 Pakistani, 2 Tunisians, 1 Turkish, and 1 Thai students.

The participants were grouped since the sample size was too small. The groups are the following: Russians, Spanish, Hungarians, Asians (including Chinese and Thai people), and African/Middle Easterners. Then, three tests had to be completed. The results of the tests were averaged to understand the participant's memory colors. Each test had a gray background with the CIELAB color code of (67.1, 0.0, 0.0).

In the first test, called "Whole palette with an image cue", many objects that are regularly seen in daily life were depicted in white with black outlines. A color picker had to be used to color these objects. Participants are free to choose any color from the color pallet. Therefore, all possible colors could be chosen. Four images had to be colored with this method. The images that need to be colored contain small details that are not necessary for the research and can be left uncolored. The areas of each image that must be colored were specified by the researcher in the room, and that part of the image had to be colored. There were no restrictions on the colors and there was no time limit. The name of the second test was "Reduced palette with an image cue". Since the test's objective was to colorize the identical images, it is comparable to the first one. The participants were unable to mix the colors in this case though. Instead, they were limited to 12 color options and they had to select from them. Thus, predefined colors were used here. They consisted of the long-term memory colors defined in other studies. Their ΔE^*_{ab} values were also considered, and the differences should be noticeable but not excessive. Prior to the testing, ten students participated in a pre-testing session where they were only re-quired to complete the coloring task using the full palette. Each memory color was averaged as a result, and these colors were added to the other 12 colors. The name of the third test was "Reduced palette without an image cue". Because the presence of an image cue can affect memory colors, this test was also conducted. There is an image cue if the object can be seen; otherwise, the participants must rely only on their imagination. Six questions were posed to the user in this case. An answer can be selected for each question from a group of 12 colors that were mentioned in the previous test. The arrow keys could be used by the participants to navigate between the color groups. The other color group does not appear instantly after pressing the forward or back arrow keys. Only the gray background is visible for three seconds. This is important to restore the participants' ability to distinguish between the two-color groups. The three tests can be seen in Figs. 3-5.

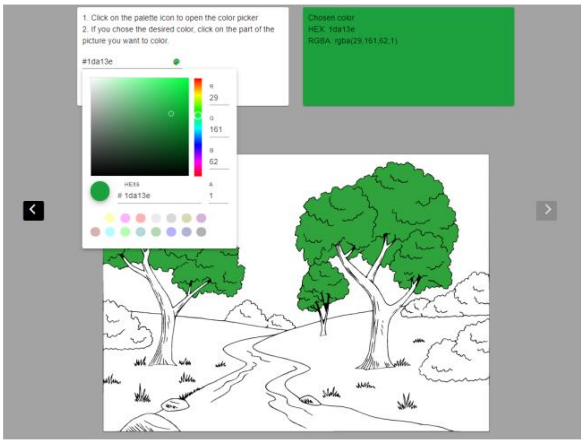


Fig. 3. Test #1: Whole palette with an image cue.

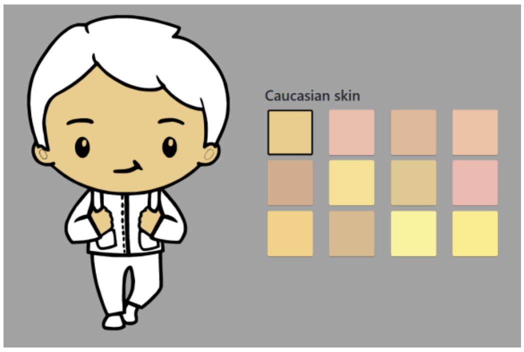


Fig. 4. Test #2: Reduced palette with an image cue.

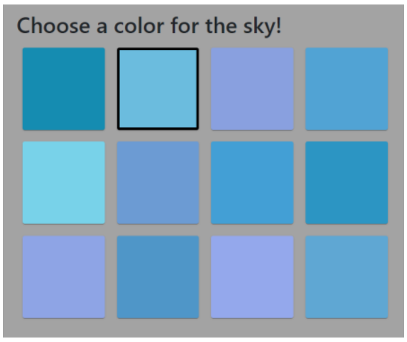


Fig. 5. Test #3: Reduced palette without an image cue.

Microsoft Excel was used to calculate the means and standard deviations. It was also used to evaluate the tests. First, for each participant and memory color examined, the mean of the three tests was determined. Additionally, these results were averaged after being divided into groups based on nationality and the duration of time spent playing virtual reality games. Then, the standard deviations for the a^* and b^* scores were determined after these groups had been formed. To understand the differences between colors, we used equation (1). The numbers in the subscript represent two different colors.

$$\Delta E^*_{ab} = \sqrt{(L^*_1 - L^*_2)^2 + (a^*_1 - a^*_2)^2 + (b^*_1 - b^*_2)^2} \quad (1)$$

Afterward, the level of difference could be easily understood based on the values presented in Table I.

TABLE I
DISTANCES BETWEEN COLORS.

ΔE^*_{ab}	Meaning
$0 \leq \Delta E^*_{ab} \leq 0.5$	No, or barely noticeable difference
$0.5 \leq \Delta E^*_{ab} \leq 1.5$	Barely noticeable difference
$1.5 \leq \Delta E^*_{ab} \leq 3$	Noticeable difference
$3 \leq \Delta E^*_{ab} \leq 6$	Obvious difference
$6 \leq \Delta E^*_{ab}$	Large difference

Lastly, comparison was made. We compared our results between our groups. Our results were also compared to the ones of Bartleson [6], Tarczali [26], Sik-Lanyi [25], and to results received in our own preliminary survey.

IV. RESULTS

The results are divided into three subsections. In the first, they are grouped by nationality, while they are grouped by virtual reality video game playtime in the second one. The importance of an image cue is investigated in the third one. In the next subsections, the following abbreviations are used: Hu (Hungarians), As (Asians), Spanish (Sp), Russians (Ru), and African/Middle-Easterners (A/ME).

A. *Results grouped by nationality*

First, the memory color of skin was assessed. It can be observed in Fig. 6 that the Asian group remembered this color differently from the other ones.

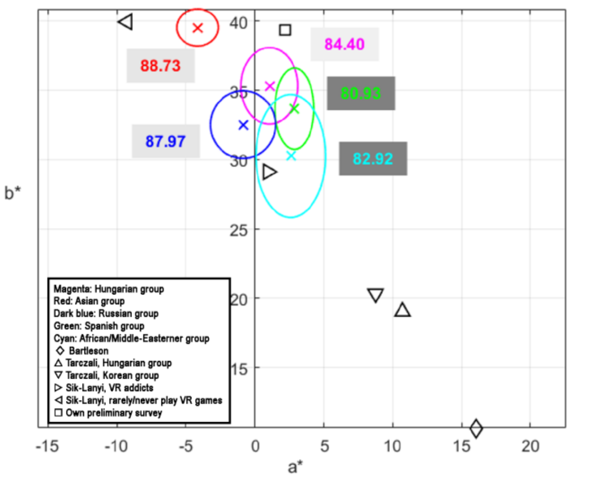


Fig. 6. Results obtained for the memory color of skin color in the CIELAB system.

The results of the remaining groups were more similar to each other, but the differences in lightness values were more noticeable. Among the four groups, Russian participants had the brightest memory color, while the Spanish group had the darkest one. These results differ the most from those that are found in Bartleson's research. The memory color of the Asian group was most similar to that of students who rarely or never play VR games in Sik-Lanyi's study. Compared to this research as well, the memory color of the Russian group was most similar to that of students who spend lots of time with computer games. Since Tarczali's research examined Far Eastern complexions in the case of Korean testers and Caucasian complexions in the case of Hungarian testers, it only makes sense to compare these results with our Hungarian group. More yellowish and fewer red results were obtained for each of these nationalities, as shown by the positive shift on the b^* axis and the negative shift on the a^* axis.

In case of the orange color, there is smaller variation between the groups as opposed to skin color. This is confirmed by the size of the ellipses as can be observed in Fig. 7. The difference between some groups, for example between the Spanish and Asian groups, was barely noticeable ($\Delta E_{ab}^* < 1.5$). The largest difference was also only 5.58. Compared to skin tone, it was 13.02. There were no outlier groups for this memory color. In case of lightness value, there was only a difference of 2.86 between the lightest and darkest. The memory color of the Asian group was slightly brighter than that of the Hungarian group. Similar results were obtained in Tarczali's research, but the other group only consisted of Korean participants in it.

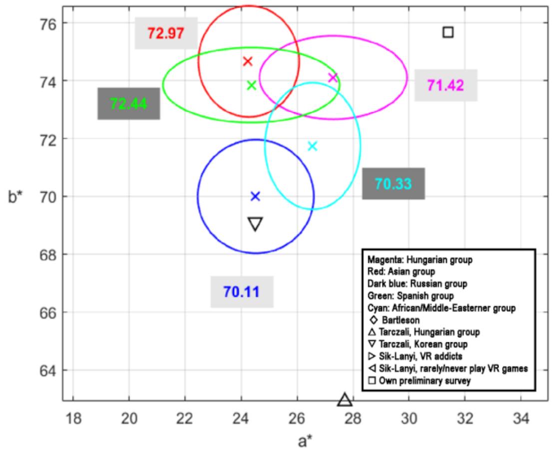


Fig. 7. Results obtained for the memory color of orange in the CIELAB system.

The next to investigate was the memory color of banana. The results are shown in Fig. 8.

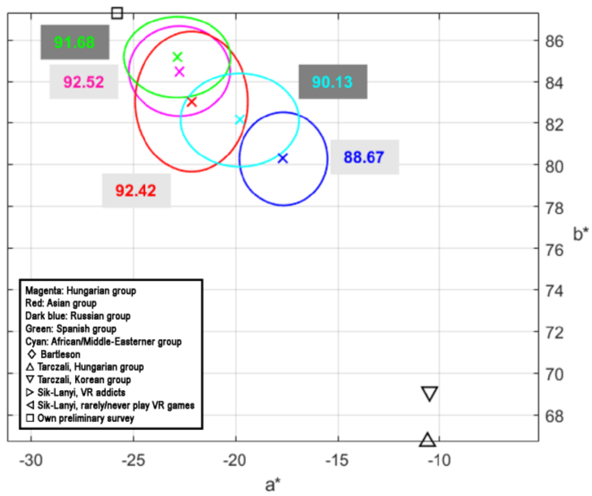


Fig. 8. Results obtained for the memory color of banana in the CIELAB system.

There was no large difference between the groups: the Hungarian group chose the brightest value, while the Russian group the darkest one. The results were not much different from the preliminary testing with 10 students, but they were clearly different from the results in Tarczali's study. In that research, it was concluded found that Korean students remembered brighter than Hungarian students. In our case however, there was no difference found between the Asian and Hungarian groups in terms of the lightness value of the banana.

Next, the memory color of the sky was investigated. The results of this examination can be observed Fig. 9.

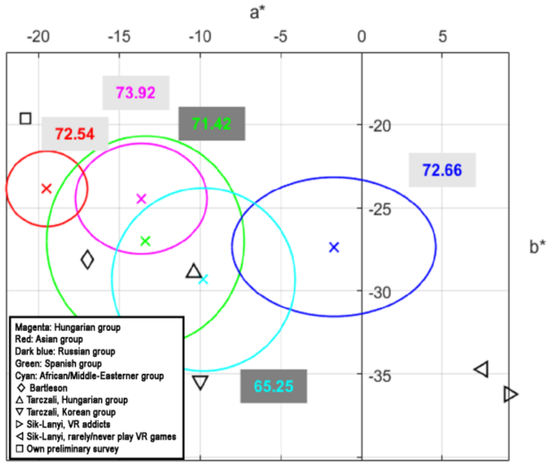


Fig. 9. The results obtained for the memory color of sky in the CIELAB system.

There were greater differences in this case, both in terms of hue and lightness. The African/Middle-Easterner group had the darkest memory color of the sky, while the Hungarian group had the lightest one. The Russian group had the most different results in terms of hue. They showed a large deviation from all groups ($\Delta E_{ab}* > 12$). In Tarczali's study, the memory colors of Korean and Hungarian skies were similarly bright, and this fact was also true in our case for the Asian and Hungarian groups.

The memory color of leaf was investigated next, and the results can be observed in Fig. 10.

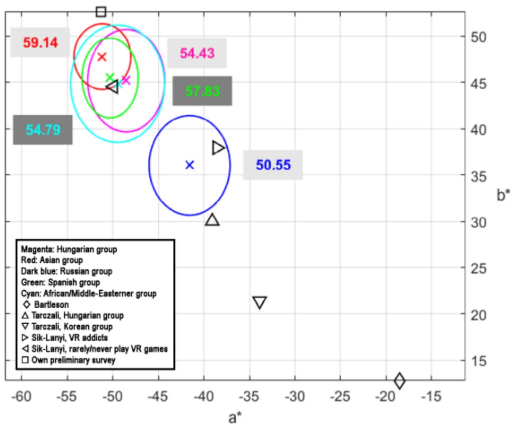


Fig. 10. The results obtained for the memory color of leaf in the CIELAB system

In this case, the memory color of the Russian group was the most different in terms of hue, and had the darkest color. The other four groups differed less from each other in hue. There was only a noticeable difference in lightness. The Asian group has the brightest memory color. Compared to the literature, four similar groups had the most similar results to those students who play VR games less often, whereas the Russian group's result was similar to the results of students who play with VR games often. Tarczali concluded in her research that the Korean leaf was less saturated compared to the Hungarian one. In our case there was no large difference between the two groups.

The memory color of grass was investigated next and the results can be observed in Fig. 11.

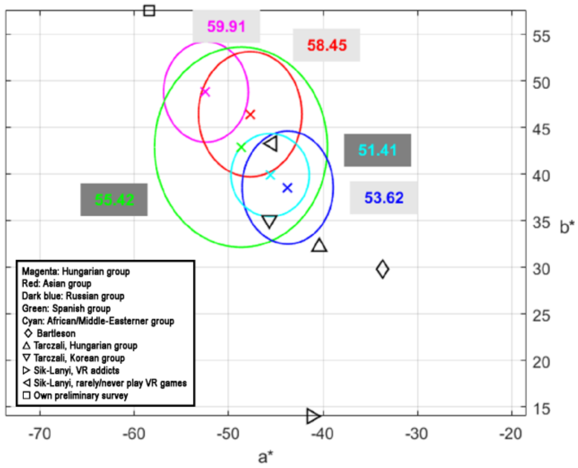


Fig. 11. Results obtained for the memory color of grass in the CIELAB system.

In contrast to the memory colors of leaf, none of the groups had outlier results. The Hungarian group produced the brightest memory color, the other ones the darkest. For leaf, the results of the Russian group were very similar to that of students who played several hours of video games in Sik-Lanyi's study, but this case was not true for the grass. The results do not differ from the Bartleson study to the same extent as for skin color or leaf. In Tarczali's research, the memory color regarding Korean grass was brighter and more saturated than that of the Hungarian grass. However, in our research, opposite results were found. The Asian group yielded a less saturated color than the Hungarian one. The results were only slightly darker.

Lastly, the memory color of river was examined. The results are shown in Fig. 12.

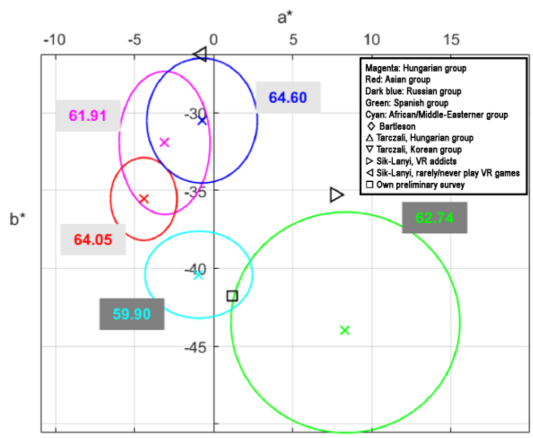


Fig. 12. The results obtained for the memory color of river in the CIELAB system.

For this memory color, the Russian group chose the brightest colors and the other ones the darkest colors. There were also a variety of results regarding hue, since none of the comparisons were below the “obvious difference” category. The colors of the Hungarian, Asian and Russian groups were closer to the memory colors of the students who rarely played VR games in Sik-Lanyi’s study. The results of Spanish and African/Middle-Easterner groups were closer to the memory colors of students who play several hours of VR games in the same research.

B. Results grouped by the amount of time spent playing virtual reality games

Next, the participants were divided into two groups based on their virtual reality game time: these groups are students who play many hours of VR games (at least 20 hours a week) or not.

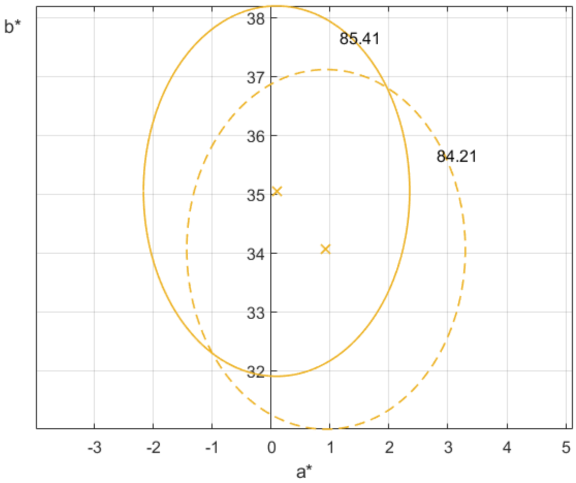


Fig. 13. Results obtained for the memory color of skin color and their dispersion ellipses. Solid line: participants who played few hours, dashed line: participants who played several hours.

Fig. 13 shows the results of students who play few hours (solid line) and of those who play several hours (dashed line) of these types of games. For the memory color of skin, the difference between the two groups was small ($\Delta E_{ab}* = 1.39$). In the study of Sik-Lanyi, there was a significant difference ($\Delta E_{ab}* = 20.17$) between the results of the same two groups.

Fig. 14 shows a comparison of the memory colors of oranges and bananas in separate coordinate systems.

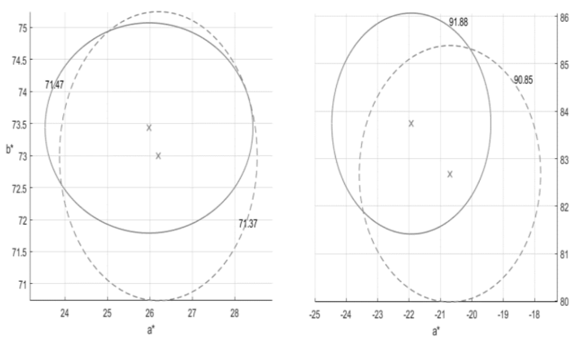


Fig. 14. Results and dispersion ellipses regarding the memory colors of oranges (left) and bananas (right). Solid line: participants who played few hours, dashed line: participants who played several hours.

In case of these two memory colors, there was no difference between participants who play several or few hours. The difference for bananas falls into the "barely noticeable" category ($\Delta E_{ab}* = 1.68$), while the difference for oranges falls into the “not, or barely noticeable” category ($\Delta E_{ab}* = 0.18$).

Fig. 15 shows the results of the memory colors of river and sky in a shared coordinate system.

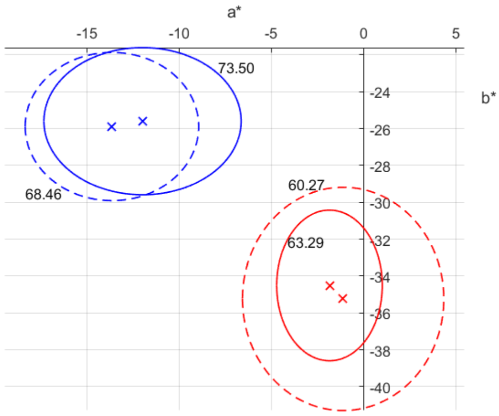


Fig. 15. Results and dispersion ellipses for the memory colors of sky (dark blue) and river (red). Solid line: participants who played few hours, dashed line: participants who played several hours.

We can already see differences in hue and lightness between students who played several or few hours. In case of both memory colors, the colors of those who played several hours were slightly darker. The difference between the two groups were in the category of “obvious difference” (sky: $\Delta E_{ab}* = 5.57$; river: $\Delta E_{ab}* = 3.11$). In Sik-Lanyi’s study, there was a similar magnitude of difference between the two groups for the sky, but the memory color of students who play several hours was not darker, whereas it was is darker in our case. In the previously referenced study, there was a large difference between the colors of river. However, it was not as large as in our case, but those who play several hours remember darker rivers, whereas in the other study it was just brighter.

Fig. 16 shows the memory color results for leaf and grass in a shared coordinate system.

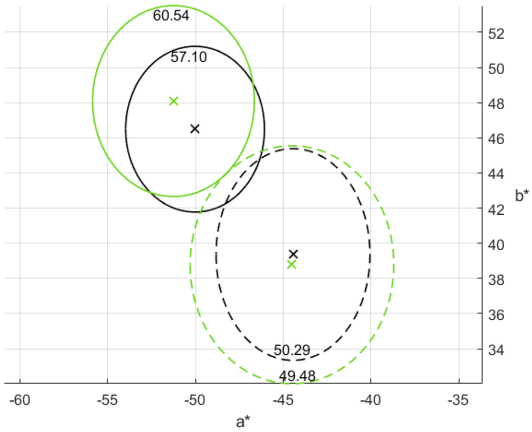


Fig. 16. Results and dispersion ellipses regarding the memory colors of leaf (black) and grass (green). Solid line: participants who played few hours, dashed line: participants who played several hours.

For these two memory colors, the difference between the two groups was most noticeable. For leaf, large differences were found ($\Delta E_{ab}* = 11.61$), while for grass the value was even greater ($\Delta E_{ab}* = 15.99$). A surprising result was that both groups chose similar colors for grass and leaf. In case of both memory colors, the darker value was obtained by the group who played several hours. These results strengthen the conclusions of Sik-Lanyi, although the two groups do not differ to the same extent. In the referenced study, the differences were even larger (leaf: $\Delta E_{ab}* = 24.03$; grass: $\Delta E_{ab}* = 54.31$).

C. *Results grouped by whether an image cue is present*

As was already described in the study, 54 people participated in the study regarding memory colors. They performed three tests for seven memory colors. In both cases, the colors for the second and third tasks were the same. However, their order was different. Out of all choices (378), 110 were the same and 268 were different. Thus, despite choosing from the same color palettes, different colors were selected in 70.90% of the decisions. This confirms the fact that the presence of the image cue influences the selection of colors.

V. Discussion

Regarding memory colors, banana and orange were those two memory colors that the students agreed on the most. Blue and green memory colors, on the other hand, led to smaller agreements between students. We can also conclude that memory colors can occasionally differ by nationality. It should be noted that not all memory colors differ, and not across all nationalities. The findings in the literature can be extended by our results, since culture can influence how colors are remembered [28]. If the results were grouped by virtual reality game playtime, a similar effect could be seen. Since not all memory colors are affected, it can only be partially verified whether virtual reality games have an impact on long-term memory colors. The memory colors of some objects were darker for people who spent more than 20 hours a week playing virtual reality games. These results build on the literature since playing video games caused grass to appear darker in people's memories [24, 25]. The choice of colors was similarly influenced by the presence of an image cue, which also supports the findings in the literature [27].

When compared to the results reported in the research of Bartleson [6], the memory colors of leaf, skin, and grass greatly differ from the findings of this study. However, this difference is presumably due to the fact that our testing was carried out on a laptop.

Naturally, this research has its own limitations as well. The tests should be conducted with additional people to receive more accurate results. Additionally, as can be seen, numerous international groups' results and that of the students from Hungary were similar. The following question might be asked given that the majority of students had spent more time in Hungary: Can a longer stay in a foreign county affect memory long-term memory colors? Further investigation is necessary in this regard.

VI. Conclusion

To investigate the cognitive effect of what we see, an application was developed during this study. It used the CIELAB color space. Three types of tests were conducted regarding memory colors.

It could be observed that the dispersion ellipses of banana and orange were the smallest. These two colors had the highest agreement between the students. The memory colors of river and grass produced the largest dispersion. Differences between memory colors of various nationalities were found, but not for all memory colors, and not between all nationalities. Playing virtual reality games also influenced memory colors, although this was also partially confirmed, since this cannot be said for all memory colors. Regarding image cues, their presence had an effect on memory colors.

The colors of the second and third tasks were the same, just mixed. In 70.90% of the decisions, a different color was chosen in the third task, despite the fact that the subjects could choose from the same set of colors. This confirms the effect of the presence or absence of an image cue on color choice.

Acknowledgment

This work has been implemented by the TKP2021-NVA-10 project with the support provided by the Ministry of Culture and Innovation of Hungary from the National Research, Development and Innovation Fund, financed under the 2021 Thematic Excellence Programme funding scheme. We would like to thank Anna Dömötör for developing as well as organizing the tests.

References

[1] Sik-Lanyi, C. Testing methods of Colour-Fidelity and Barrier-Free Design of Virtual Worlds. D. Sc. Theses, Hungarian Academy of Sciences, Budapest, Hungary, 2020

[2] Natural Color System – The International Language of Color Communication. Available online (accessed on 12.02.2023.): http://www.ncscolour.com/

[3] MacDonald, L. *Colour in Computer Graphics*; Lecture Notes, 2nd edition, MacColor Ltd, 1998.

[4] Elgendy, M. Indoor Navigation for People with Visual Impairment. PhD Thesis, University of Pannonia, Veszprem, Hungary 2021

[5] Hering, E. *Outlines of a Theory of the Light Sense*; Harvard University Press: Cambridge, MA, 1964; ISBN 9780674649002

[6] Bartleson, C. J. Memory Colors of Familiar Objects. *J. Opt. Soc. Am.* 1960, 50, 73–77, DOI: 10.1364/josa.50.000073

[7] Jin, E. W.; Shevell, S. K. Color Memory and Color Constancy. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 1996, 13, 1981–1991, DOI: 10.1364/josaa.13.001981

[8] Siple, P.; Springer, R. M. Memory and Preference for the Colors of Objects. *Percept. Psychophys.* 1983, 34, 363–370, DOI: 10.3758/bf03203049

[9] Nilsson, T. H.; Nelson, T. M. Delayed Monochromatic Hue Matches Indicate Characteristics of Visual Memory. *J. Exp. Psychol. Hum. Percept. Perform.* 1981, 7, 141–150, DOI: 10.1037//0096-1523.7.1.141

[10] Newhall, S. M.; Burnham, R. W.; Clark, J. R. Comparison of Successive with Simultaneous Color Matching. *J. Opt. Soc. Am.* 1957, 47, 43, DOI: 10.1364/josa.47.000043

[11] Bodrogi, P. Shifts of short-term colour memory. PhD Thesis, University of Pannonia, Veszprem, Hungary, 1998.

[12] Baranyi, P.; Csapó, Á. Definition and synergies of cognitive infocommunications. *Acta Polytech. Hung.* 2012, 9, 67–83.

[13] Baranyi, P.; Csapo, A.; Sallai, G. *Cognitive Infocommunications (CogInfoCom)*; Springer International Publishing: Cham, Switzerland, 2016; ISBN 9783319359472

[14] Baranyi, P.; Csapo, A.; Varlaki, P. An Overview of Research Trends in CogInfoCom. In *Proceedings of the IEEE 18th International Conference on Intelligent Engineering Systems INES 2014*; IEEE, 2014.

[15] Sallai, G. The cradle of cognitive infocommunications. *Acta Polytech. Hung.* 2012, 9, 171–181.

[16] Macik, M. Cognitive Aspects of Spatial Orientation. *Acta Polytech. Hung.* 2018, 15, 149–167.

[17] Kovari, A.; Katona, J.; Costescu, C. Quantitative Analysis of Relationship between Visual Attention and Eye-Hand Coordination. *Acta Polytech. Hung.* 2020, 17, 77–95, DOI: 10.12700/aph.17.2.2020.2.5

[18] Katona, J. Measuring Cognition Load Using Eye-Tracking Parameters Based on Algorithm Description Tools. *Sensors (Basel)* 2022, 22, 912, DOI: 10.3390/s22030912

[19] Katona, J. Clean and Dirty Code Comprehension by Eye-Tracking Based Evaluation Using GP3 Eye Tracker. *Acta Polytech. Hung.* 2021, 18, 79– 99, DOI: 10.12700/aph.18.1.2021.1.6

[20] Sudar, A.; Csapo, A. B. Interaction Patterns of Spatial Navigation in VR Workspaces. In *Proceedings of the 2019 10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*; IEEE, 2019, DOI: 10.1109/CogInfoCom47531.2019.9089998

[21] Torok, A. From Human-Computer Interaction to Cognitive Infocommunications: A Cognitive Science Perspective. In *Proceedings of the 2016 7th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*; IEEE, 2016, DOI: CogInfoCom.2016.7804588

[22] Wolfe, J. M.; Kluender, K. R.; Levi, D. M. Sensation & Perception; Oxford University Press, 2018

[23] Hansen, T.; Olkkonen, M.; Walter, S.; Gegenfurtner, K. R. Memory Modulates Color Appearance. *Nat. Neurosci.* 2006, 9, 1367–1368, DOI: 10.1038/nn1794

[24] Sik-Lányi, C.; Sik, A.; Sik, G. "Virtual world – „A brave new world" – but what kind of colours are used in these virtual environments?" Poster, 26TH Session of the CIE, 4 July – 11 July 2007, Beijing, China, vol. 1. D1, 99–102, 2007

[25] Sik-Lányi, C. Investigating of Memory–Colours of Intellectually Disabled Children and Virtual Game Addict Students. In *Symposium of the Austrian HCI and Usability Engineering Group*; Springer: Berlin, Heidelberg, 2009; pp. 463–475

[26] Tarczali, T.; Park, D.-S.; Bodrogi, P.; Kim, C.Y. Long-Term Memory Colors of Korean and Hungarian Observers. *Color Res. Appl.* 2006, 31, 176–183, DOI: 10.1002/col.20192

[27] Tarczali, T. Investigation of Colour Memory. PhD Thesis, University of Pannonia, Veszprem, Hungary, 2007

[28] Smet, K. A. G.; Lin, Y.; Nagy, B. V.; Németh, Z.; Duque-Chica, G. L.; Quintero, J.M.; Chen, H.-S.; Luo, R. M.; Safi, M.; Hanselaer, P. Cross-Cultural Variation of Memory Colors of Familiar Objects. *Opt. Express* 2014, 22, 32308–32328, DOI: 10.1364/OE.22.032308

[29] Sik-Lanyi, C.; Miklós, B.; Szőcs-Józsa, T. Well-Known Brands for Testing Colour Memory. In Pannonian Conference on Advances in Information Technology (PCIT 2020); 48–55, 2020

[30] Hill, B.; Roger, T.; Vorhagen, F. W. Comparative Analysis of the Quantization of Color Spaces on the Basis of the CIELAB Color-Difference Formula. *ACM Trans. Graph.* 1997, 16, 109–154, DOI: 10.1145/248210.248212

Cecilia Sik-Lanyi studied Mathematics and Computer Science. She obtained the degree dr.univ (1993) and of Ph.D. (2000) at the University of Pannonia, D.Sc. (2022) at the Hungarian Academy of Science. Currently she is a professor at the University of Pannonia. Her research area is Virtual Reality, Serious Games and Colour Science. She has published more than 500 referred articles and conference papers and worked as guest editor for many renewed journals and books. She received several awards, the most important ones are:

– "Master teacher" award of the Hungarian Ministry of Education in 2001,
– "AAATE-diamond-award" of the Association for the Advancement of Assistive Technology in Europe in 2015,
– "King Salman Award for Disability Research" of the King Salman Center for Disability Research in the Kingdom of Saudi Arabia in 2018.

Bence Halmosi is a PhD student at the Faculty of Information Technology of the University of Pannonia, Hungary. His research area include artificial intelligence, human-computer interaction and virtual reality.

Jinat Ara graduated from Southeast University, Bangladesh in 2018 with a bachelor's degree in Computer Science and Engineering (CSE). Following that she obtained her master's degree in 2020 from Jahangirnagar University, Bangladesh from the Department of Computer Science (CS). At the same time, she worked as a Software Engineer at Odyssey Apps, Dhaka, Bangladesh. Currently, she is pursuing her Ph.D. degree at the Faculty of Information Technology, University of Pannonia, Hungary. Her research area for her Ph.D. studies on the area of "accessibility investigation in web platforms or environments". Additionally, she has a keen interest in conducting cutting-edge research and currently, she is carrying around six years of research experience in several areas of computer science. Her area of research interest is 'Machine Learning (ML)', 'Virtual and Augmented Reality', 'Assistive Technologies', 'Web Technologies', 'Artificial Intelligent (AI)', 'Natural Language Processing (NLP)' and 'Automated Vehicles (AV)'.

Judit Szűcs is an Assistant Professor at the Faculty of Information Technology of the University of Pannonia, Hungary. She received her PhD degree in 2021 in Computer Science at the University of Szeged. Her main interests include binary tomography, image processing, and computer vision. She has 30 scientific publications which include 6 journal papers, and 19 studies found in various conference proceedings. She is an IEEE member since 2023.

Tibor Guzsvinecz is an Assistant Professor at the Faculty of Information Technology of the University of Pannonia, Hungary. He obtained his PhD degree in the field of Information Science in 2021. His research areas are data science, gamification, human-computer interaction, spatial ability, and virtual reality. He authored 69 scientific publications which include 19 journal papers, and 41 studies found in various conference proceedings. At the moment, he is an editor of a book project about gamification. He is an IEEE member since 2023.

The Accuracy of the k-Nearest Neighbors and k-Means Algorithms in Gesture Identification

Tibor Guzsvinecz, Judit Szűcs, Veronika Szucs, Robert Demeter, Jozsef Katona and Attila Kovari

Abstract—**In today's digital era, human-computer interaction interfaces evolve and increase together with the needs of the users. However, the existing technologies have their limitations, which can hinder the efficiency of modern input devices like the Kinect sensor or other similar sensors. In this paper we improved our previous algorithm by extending it with two algorithms that aim to help telerehabilitation for individuals with movement disabilities. These two algorithms are based on the k-Nearest Neighbors, the k-Means algorithms. The algorithms are designed to accommodate the needs of the patients by adapting to their gestures based on their previous three. Using these gestures, the algorithms create multiple gesture acceptance domains around each coordinate of the gesture. Consequently, they decide whether the next user-input gesture can be considered the same movement. The accuracy of these algorithms was evaluated in three acceptance domains by comparing gesture descriptors with either the Euclidean or the Manhattan distance calculation methods. The results show that k-Nearest Neighbors algorithm yields better results in larger acceptance domains, while the k-Means algorithm can provide a better gesture acceptance rate in the smaller ones. The results show that both algorithms can be used in the telerehabilitation process, although the k-Means algorithm is more accurate than the k-Nearest Neighbors algorithm.**

Index Terms—**cognitive aspects of virtual reality, cognitive infocommunications, human-computer interaction, Kinect, motivation, real-time gesture recognition, rehabilitation**

I. INTRODUCTION

In a virtual reality (VR) system, the human plays a crucial role as a key element [1]. The interaction, usability, and comfort between the human and the machine are essential factors to consider. These aspects are addressed in the field of Human-Computer Interaction (HCI), which is a multidisciplinary research area focused on studying and addressing such questions [2]. The fields of Cognitive Aspects of Virtual Reality (cVR) and Cognitive InfoCommunications (CogInfoCom) both address HCI, and one of their main focuses is human cognition. They both focus on showcasing the latest advancements in information and communication technologies (ICT) that facilitate the interaction between humans and machines [3–8]. Their objective is to enhance, restore, or even develop new cognitive abilities in users by utilizing ICT engineering tools and model-based approaches. Not to mention, CogInfoCom and cVR are also closely related [9].

Fortunately, both fields also focus on human-machine blended interaction, particularly in the context of evaluating gestures and movements [10–18]. This opens up new ways for analyzing several human factors using novel cognitive IT approaches. These fields of research also delve into various areas including eye-tracking [19], brain-computer interfaces (BCIs) [20, 21], VR systems [22, 23], virtual laboratories [24, 25], gamification [26], sentiment analysis [27], and other educational environments [28, 29]. The findings from these studies hold significant value in fields such as education, development, and rehabilitation.

Motion rehabilitation stands out as a crucial and significant application area, particularly in the field of healthcare information technology. This area of research is quite important since many people suffer from physical disabilities. One of their causes is stroke, which is a frequent disease in modern society, and has a great impact on the human population. Research studies have shown that 48% of individuals who have survived brain-to-asthma disease suffer from half-side paralysis [30, 31]. Additionally, cognitive decline can be detected in over 60% of cases. Aphasia, a language impairment, affects approximately 12-18% of stroke patients. Furthermore, 24-53% of individuals with stroke become partially or completely dependent on others for their daily activities. Given these challenges, it is imperative to incorporate modern technology to address the needs of these patients [32, 33].

Therefore, rehabilitation for patients can be made more interesting by incorporating engaging and modern environments for exercises or gesture tracking. Many computer-aided solutions are already available that assist with traditional therapy. However, these solutions often require direct interaction with a computer, such as pressing keys or using a mouse, which can pose challenges for some patients. Bruno et al. explored the potential of using commercial video games in rehabilitation. They examined 4,728 relevant articles

Submitted on 6/13/2023

Tibor Guzsvinecz is with the Department of Information Technology and its Applications, Faculty of Information Technology, University of Pannonia, Zalaegerszeg, Hungary (e-mail: guzsvinecz.tibor@zek.uni-pannon.hu). He is the corresponding author.

Judit Szűcs is with the Department of Information Technology and its Applications, Faculty of Information Technology, University of Pannonia, Zalaegerszeg, Hungary (e-mail: szucs.judit@zek.uni-pannon.hu).

Veronika Szucs is with the Department of Electrical Engineering and Information Systems, Faculty of Information Technology, University of Pannonia, Veszprem, Hungary (e-mail: szucs.veronika@mik.uni-pannon.hu).

Robert Demeter is with the Department of Automation and Information Technology, Transilvania University of Brasov, Romania (e-mail: rdemeter@unitbv.ro).

Jozsef Katona is with the Department of Software Development, Institute of Computer Engineering, University of Dunaujvaros, Dunaujvaros, Hungary; and Kandó Kálmán Faculty of Electrical Engineering, Óbuda University, Budapest, Hungary (e-mail: katonaj@uniduna.hu).

Attila Kovari is with the Institute of Digital Technology, Faculty of Computer Science, Eszterházy Károly Catholic University, Eger, Hungary; Kandó Kálmán Faculty of Electrical Engineering, Óbuda University, Budapest, Hungary; Department of Natural Sciences, Institute of Engineering, University of Dunaújváros, Dunaujvaros, Hungary; and GAMF Faculty of Engineering and Computer Science, John V on Neumann University, Kecskemét, Hungary (e-mail: kovari.attila@uni-eszterhazy.hu).

DOI: 10.36244/ICJ.2024.5.4

yield results that are comparable to conventional rehabilitation methods [34, 35]. Still, it is argued by Ghazarian and Noorhosseini that exercise games may not meet the needs of patients if the application relies solely on pre-calibrated settings or predetermined correction values [36].

Extracting motion descriptors from the sensors that are used as the input interface is a crucial aspect of controlling applications with motion. Our previous work aimed to retrieve, process, analyze, and improve the motion descriptors that were obtained from the Microsoft Kinect sensor [17, 18]. In the previously referenced studies, we presented multiple algorithms that are easy to use in a home environment as a part of telerehabilitation. Certain gestures can be recommended by the therapist to the patients and they can exercise in their homes. This way, they can also receive help from their families if needed. Based on the results of the exercises, it can be decided whether the patient made any progress during the physical rehabilitation process. Low-cost sensors, such as the Microsoft Kinect, are also commonly used in the telerehabilitation process [37].

As can be observed, our goal is to support telerehabilitation. Thus, this study aims to provide another set of algorithms and methods to the existing ones, thus increasing the number of algorithms to choose from for gesture recognition. For this goal, we have expanded our previous work with the k-Nearest Neighbors (k-NN) and k-Means algorithms [38, 39]. These algorithms are evaluated in this study to see how well they support gesture identification.

Therefore, this study is structured as follows. Section II provides an insight into problems regarding gesture recognition with the Microsoft Kinect as well as into our previous work. Section III presents the materials and methods used in this study. The results and discussion are detailed in section IV, while the limitations and future plans regarding this research can be observed in section V. Conclusions are drawn in section VI.

II. Problem identification and previous solution

Nowadays, hospitals are overcrowded, but this issue can be alleviated by telerehabilitation. Among others, the Microsoft Kinect sensor offers a cost-effective alternative to more expensive sensors. This makes this sensor accessible for people with movement disabilities to do physical rehabilitation exercises at their homes [37]. However, to achieve this, an easy-to-use application with algorithms that are specifically designed for the Kinect is necessary.

With the Software Development Kit (SDK) developed by Microsoft, the Kinect sensor can provide real-time x, y, z coordinates for various body joints of the user. It uses a built-in coordinate system for this process. However, the following two problems arise when the Kinect sensor creates these coordinates.

The first problem concerns the position of users, and thereby of gestures. The Kinect sensor assigns different coordinate values to the same gesture when the user stands in a different position in front of it. This also happens when the Kinect is placed in a different position. Therefore, algorithms that can sense various positions need to be developed.

The second problem concerns the speed of movement of users. The mentioned speed is measured in frames per second. Naturally, slower gestures have more frames, while faster movements have fewer frames. Similarly to the first problem, the various numbers of frames make it challenging to recognize the same gestures as the corresponding movement descriptors can be on another frame. This means that depending on the speed of the gesture, the same movement can have different number of frames. Therefore, some normalization method is required to address the varying number of frames.

To solve these problems, we originally proposed the Asynchronous Prediction-Based Movement Recognition (APBMR) algorithm in a previous study [18]. The algorithm does not require extensive computational power, supports low-cost sensors like the Microsoft Kinect, can be used during telemedicine, and provides more precise adaptation to the needs of the patients. The APBMR algorithm utilizes prediction techniques to anticipate the user's next gesture based on the preceding three movements. It then determines whether the forthcoming gesture corresponds to the same movement, with the aim of keeping the motivation of patients. For this, it creates three acceptance domains at ±0.05 m, ±0.10 m, and ±0.15 m of each coordinate of a gesture. The goal of the user is to stay inside these acceptance domains in each frame. Since all frames are evaluated, a gesture can be considered accepted if its coordinates are at least 50% inside in these acceptance domains. If one is accepted inside the strictest acceptance domain, that would mean that the APMBR algorithm can accurately predict and classify the gesture of the user based on the previous three gestures. Additionally, it monitors the patient's position and the gesture's speed.

III. Materials and methods

Both the implemented k-NN and k-Means algorithms use the same principle as the previously mentioned APBMR algorithm. In fact, its source code and graphical user interface (GUI) were modified, and these new algorithms were integrated into them. Both of the GUIs were created in WPF using C#. The modified GUI can be observed in Fig. 1.

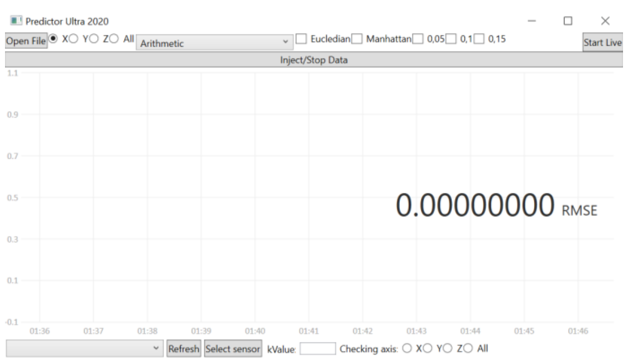


Fig. 1. The GUI of the application that is used in this study.

More checkboxes and radio buttons can be found in the newer version of the GUI. Compared to the older one, the acceptance domains of the gestures can be selected as can be seen in upper right corner in Fig 1. These are ±0.05 cm, ±0.10 cm, and ±0.15 cm. Originally, the APBMR algorithm evaluated the gestures in all three, but here, it is possible to evaluate them separately. Also, in the original version, Euclidean distances were compared between the older gestures and the forthcoming ones. Now it is possible to compare Manhattan distances as well. There is also a new option: The gestures can now be displayed on certain axes. Any axis can be chosen by itself, and there is an option to display the gesture descriptors on all axes simultaneously.

As was mentioned, the application was extended with the k-NN and k-Means algorithms. These algorithms can only work when all three axes are investigated. This means that they cannot handle coordinates separately for each axis. These implemented algorithms require a new parameter called k. This can be chosen by the user, but it must be an integer. Based on empirical testing, the value of k should be between 1 and 10. These algorithms are elaborated on in the following subsections.

Similarly to the APBMR algorithm, the gesture recognition process with both the k-NN and k-Means algorithms used the previous three movements of the user to create acceptance domains for the next gesture. This was done so that the algorithm could easily adapt to any change in the user's next movement. If we increased or decreased the number of used gestures, the algorithms became less accurate. Changing the value of k did not affect the adaptability of the algorithms. However, by increasing the value of k the computational time was also increased.

Both algorithms were evaluated using three different gestures (triangle, infinity symbol, waving) that were repeated ten times. The movement descriptors of these gestures were imported from a file that contained recorded coordinates of these gestures. We used a recording that was created in 2019 with 48 students. 33 were male and 15 were female. They were 22.3 years old on average with a standard deviation of 2.8 years. Their average height was 178.1 cm with a standard deviation of 10.2 cm. There were no selection criteria to join the research, thus every person who was willing could help with the measurements. Each movement was done 10 times by one participant and they were asked to slowly change position during the process while repeating the gestures. However, they always had to face the Microsoft Kinect sensor which captured each frame of the gestures and gave them x, y, z coordinates in its own coordinate system.

The accuracy of these two algorithms was evaluated in all three acceptance domains (±0.05 cm, ±0.10 cm, and ±0.15 cm) with two distance calculation methods (Euclidean and Manhattan). A gesture was considered acceptable if at least 50% of it was in the chosen acceptance domain.

A. *The k-NN algorithm*

Essentially, this algorithm assumes that other coordinates are located in close proximity to the coordinate that is investigated at a moment. Once the value of k is determined, it looks for the k nearest neighbors to the coordinate that is investigated. The average of these distances provides an approximate estimation. The steps of the algorithm are the following in the application.

1. The data to be examined have to be loaded.
2. A value for integer k has to be chosen.
3. Iterates through all elements in the loaded dataset.
4. Calculates the distances or differences between each coordinate and the one that is investigated at the moment.
5. Sorts the calculated distances in ascending order.
6. Selects the k smallest distances after sorting.
7. Calculates the averages of these k distances.

As can be seen, no significant modifications were made to this algorithm, and its basic principle remained largely intact. However, there are some differences in the implementation.

Since the estimation of a forthcoming gesture is based on the previous three, the entire dataset did not have to be considered. Instead, the coordinates of the three gestures were taken into account. Additionally, the coordinates of each axis were split into separate parts. When considering all three axes together, the results were not satisfactory. Therefore, the calculation of distances between the investigated gesture descriptors is performed on the same respective axis.

The value of k cannot be too large since there is an upper limit on the number of coordinates for each gesture. Therefore, the value of k must be smaller than the number of distances calculated for estimating a single gesture.

The obtained k distances do not explicitly indicate which coordinate they were calculated for, and no indices were saved during the examination process. Therefore, the coordinate represented by the given distance was determined based on the examined value, the obtained distance, and other data. This retrieval process uses the same distance calculation method as the one used for calculating the distances.

We also had to ensure that the distances are never zero since the algorithm iterates through all coordinates of the three gestures and calculates the distance relative to the examined gesture. It is possible for the distance to be zero at least once because it may occur between the coordinates of the three gestures. However, a zero value is not acceptable in this case because the distances are examined and the k smallest distances are selected. Thus, one or more zero values can significantly affect the estimations.

B. *The k-Means algorithm*

The k-Means algorithm is primarily used for solving clustering problems. Its principle is to analyze a dataset and approximate one or more centroids that represent the groups or clusters of data points. Contrary to the k-NN algorithm, more modifications were made to this algorithm. Its steps are the

following in the application:
1. The data to be examined have to be loaded.
2. The value of integer k has to be determined.
3. Randomly selects k data points from the dataset as initial centroids.
4. Assigns each data point to the cluster whose centroid is closest to it.
5. Calculates the distances between each data point and the centroids.
6. Groups the data points into clusters based on the distances (assigns each data point to the cluster whose centroid is closest to it).
7. Calculates the new centroids by computing the average of data points within each cluster.
8. Repeats steps 5-7 until the values of centroids do not change or converge.

If the value of k is now known in advance, the algorithm should be run with different k values and an optimal solution should be found.

During the implementation process, there were cases where, due to random selection, the centroid values were the same. This more likely occurs when a certain gesture consists of fewer frames. Therefore, if a new coordinate estimation occurs and the randomly selected values are the same for multiple centroids, the random selection will continue until all of them receive different values.

Once the clusters and the points within the clusters are found, the algorithm examines the clusters to see if they actually contain points. Since the differences between the coordinates are very small, it is possible that a cluster might not contain any points. If it does contain points, the analysis continues. The three closest coordinates from the clusters compared to the examined coordinates are selected, separately for each axis. Additionally, the algorithm finds the three closest coordinates for each axis. This way, a total of 27 coordinates are chosen from the cluster. Then, once the first closest coordinate is found, its value is modified to avoid selecting it as the second closest value as well.

In the case of having nine coordinates from a cluster, the average of the three closest coordinates is calculated for each axis separately. These steps are performed for the other clusters as well until the new estimated points are yielded. The new estimated points and their coordinates are examined separately for each axis to check if they are suitable. However, the most important criterion is that the difference between the estimated coordinate of a given axis and the corresponding investigated coordinate should be smaller or equal to 0.05. This number was chosen because it represents the smallest magnitude of the acceptance domain. Due to this, the gesture will only be accepted if it is genuinely accurate.

If any of the estimated coordinates of the clusters are not suitable, the values of the coordinates that do meet the criteria will be determined as the values of the corresponding centroids, and the cycle starts again.

Considering that the initial approximation of the result is often incorrect due to random selection and averaging, we used a selected value for k in this algorithm. This value is completely independent and can be decided by the user in each case. There is also no maximum restriction on it, except that it should not be excessively large and unnecessary. The value of k is responsible for determining how many times the cycle, i.e., the approximation and estimation process, is executed. This is necessary because there may be cases where the values of the estimated coordinates do not meet the 0.05 criterion even after multiple approximation attempts.

Empirical testing shows that the larger the value of k is, the more accurate the estimated results will be. However, if k is drastically too large, there will be no noticeable improvement in the results, but it will not affect negatively the accuracy of the results either.

IV. Results and Discussion

This section is split into three subsections. In the first one, the results regarding the k-NN algorithm are shown. The results of the k-Means algorithm are presented in the second subsection, while the application of the algorithms is presented in the third.

A. Results of the k-NN algorithm

The results regarding the k-NN algorithm can be observed in Table I. The median of accepted gestures regarding the ten repeating gestures is shown in it in each acceptance domain. Every three lines correspond to a certain gesture. The first three are the triangle, the second three are the waving, and the third three are the infinity-shaped gesture.

TABLE I
THE MEDIAN OF THE ACCEPTED NUMBER OF GESTURES WITH THE K-NN ALGORITHM.

k	Euclidean			Manhattan		
	±0.05	±0.10	±0.15	±0.05	±0.10	±0.15
1	0	4	10	0	4	10
5	0	4	10	0	4	10
10	0	4	10	0	4	10
1	10	10	10	10	10	10
5	10	10	10	10	10	10
10	10	10	10	10	10	10
1	2	10	10	1	10	10
5	2	10	10	2	10	10
10	5	10	10	5	10	10

As can be observed in Table I, the triangle gestures produced the worst gesture acceptance numbers. While the infinity-shaped gestures yielded better numbers, the waving gestures proved to be most easily recognizable. The median of accepted gestures is also quite similar between the two distance calculation methods. Fig. 2 shows how the waving gesture was predicted by the k-NN algorithm. The value of integer k was 10. The blue line represents the gesture prediction, while the orange one is the actual gesture.

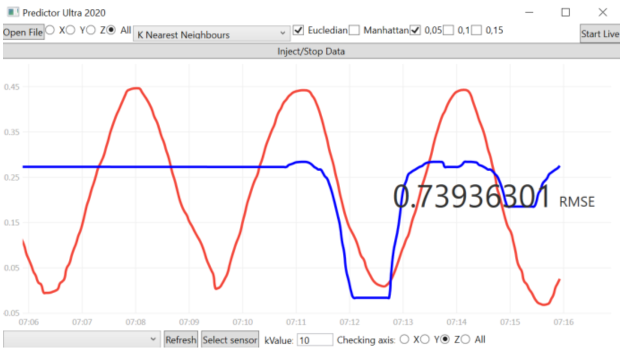


Fig. 2. Gesture prediction with the k-NN algorithm.

B. *Results of the k-Means algorithm*

Next, the results of the k-Means algorithm were investigated in a similar way. The results can be found in Table II.

TABLE II
THE MEDIAN OF THE ACCEPTED NUMBER OF GESTURES WITH THE K-MEANS ALGORITHM.

k	Euclidean			Manhattan		
	±0.05	±0.10	±0.15	±0.05	±0.10	±0.15
1	7	10	10	5	10	10
5	10	10	10	10	10	10
10	10	10	10	10	10	10
1	10	10	10	10	10	10
5	10	10	10	10	10	10
10	10	10	10	10	10	10
1	7	10	10	5	10	10
5	10	10	10	10	10	10
10	10	10	10	10	10	10

The k-Means algorithm yields better results can the k-NN algorithm. As can be seen in the case of the k-Means algorithm, the worst median of accepted gestures is five out of ten. Naturally, it is also possible to increase the number of k above 10 to enhance accuracy. However, the also increases the prediction time. Fig. 3 shows how the waving gesture was predicted by the k-Means algorithm. The parameters were the same as in Fig. 2.

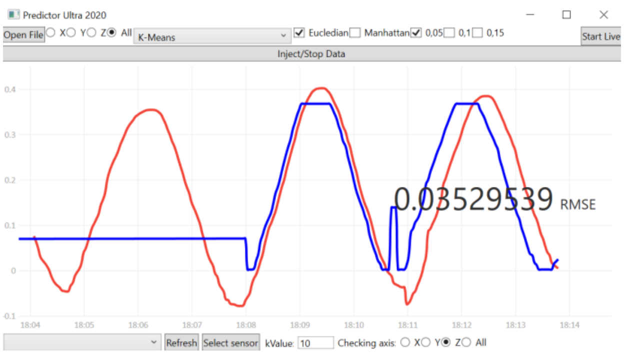


Fig. 3. Gesture prediction with the k-Means algorithm.

C. *Application of the algorithms*

The results show that both algorithms are easy-to-use and viable in recognizing the gestures of the users. Additionally, these algorithms can adapt to the current capabilities of the users provided that the best options are used. Consequently, telerehabilitation can be an option using these algorithms. Therefore, the presence of a therapist is only required during a consultation. Thus, the rehabilitation process of people with movement disabilities could be made more convenient and safer in their homes.

V. Limitations and Future Plans

As in most cases, the current application can also be improved in various ways. It may be useful to implement other algorithms (even in real-time) to provide new methods for gesture recognition. The implemented algorithms could also be run in a parallel manner, thus all of them could evaluate gestures at the same time. This would simplify the selection of the most suitable algorithm for a certain gesture. The results of the application could be stored in a database, ensuring that previous estimation results are not lost. New types of checks can also be designed to classify gestures and coordinates, resulting in multiple acceptance criteria. By further examining these results, additional conclusions could be drawn. Automating the application is also possible, thereby providing an easier-to-use user interface. In addition to the current user interface, a 3D coordinate system could be implemented. This could provide better visibility of the analyzed movements and the shape and deviations of the estimated movements.

VI. Conclusion

In this paper, an existing application was extended with two algorithms: the k-NN, and the k-Means algorithms. They were evaluated in three acceptance domains by calculating distances between gesture descriptors with either the Euclidean or the Manhattan method.

The evaluations of the algorithms clearly indicate that while the k-Means algorithm operates in a more complex manner, it is capable of providing more accurate predictions of forthcoming gestures in most cases. However, the k-NN algorithm can also prove to be more appropriate for certain gestures. In conclusion, these algorithms could be used at home, the rehabilitation process can be made easier for both the therapist and patient if they choose these methods. Based on the results however, the k-Means algorithm is more suitable for gesture recognition than the k-NN algorithm.

Acknowledgment

This work has been implemented by the TKP2021-NVA-10 project with the support provided by the Ministry of Culture and Innovation of Hungary from the National Research, Development and Innovation Fund, financed under the 2021 Thematic Excellence Programme funding scheme. We would like to thank Attila Magyar for his help when developing the previous algorithms. We would also like to thank Mate Oravecz for his help in implementing the k-NN and k-Means algorithms.

References

[1] G. Burdea and P. Coiffet, *Virtual Reality Technology*. Nashville, TN: John Wiley & Sons, 1994.

[2] J. E. Finlay, A. Dix, R. Beale, and G. D. Abowd, *Human-Computer Interaction*, 3rd ed. Philadelphia, PA: Prentice Hall, 2003.

[3] P. Baranyi and A. Csapo, "Definition and synergies of cognitive infocommunications", *Acta Polytech. Hung.*, vol. 9, pp. 67–83, 2012.

[4] G. Sallai, "The cradle of cognitive infocommunications"," *Acta Polytech. Hung.*, vol. 9, no. 1, pp. 171–181, 2012.

[5] P. Baranyi, Á. Csapó, Á. Balázs, and P. Várlaki, "An Overview of Research Trends in CogInfoCom", in *18th International Conference on Intelligent Engineering Systems – INES 2014*, Tihany, 2014, pp. 181–186.

[6] P. Baranyi and A. Csapo, *Cognitive Infocommunications*. Springer International Publishing Switzerland.

[7] A. Torok, "From human-computer interaction to cognitive infocommunications: A cognitive science perspective," in *2016 7th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, 2016, pp. 000 433–000 438.

[8] C. Vogel and A. Esposito, "Interaction Analysis and Cognitive Infocommunications," *Infocommunications J.*, vol. 12, no. 1, pp. 2–9, 2020.

[9] I. Horváth, B. Ádám, B. Csapó, A. Berki, and P. Sudár, "Definition, Background and Research Perspectives Behind 'Cognitive Aspects of Virtual Reality' (cVR)," *Infocommunications J.*, Special Issue: Internet of Digital & Cognitive realities, pp. 9–14, 2023.

[10] T. Ujbányi, A. Kővári, G. Sziládi, and J. Katona, "Examination of the eye- hand coordination related to computer mouse movement," *Infocommunications J.*, vol. 12, no. 1, pp. 26–31, 2020.

[11] A. Kovari, J. Katona, and C. Costescu, "Quantitative analysis of relationship between visual attention and eye-hand coordination," *Acta Polytech. Hung.*, vol. 17, no. 2, pp. 77–95, 2020.

[12] A. B. Csapo, H. Nagy, A. Kristjansson, and G. Wersenyi, "Evaluation of human-Myo gesture control capabilities in continuous search and select operations," in *2016 7th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, 2016, pp. 000 415–000 420.

[13] T. Guzsvinecz, V. Szucs, and A. Magyar, "Preliminary results of evaluating a prediction-based algorithm for movement pattern recognition and classification," in *2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, 2020, pp. 000039–000044.

[14] G. Sziladi, T. Ujbanyi, J. Katona, and A. Kovari, "The analysis of hand gesture based cursor position control during solve an IT related task," in *2017 8th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, 2017, pp. 000 413–000 418.

[15] V. Szucs, T. Guzsvinecz, and A. Magyar, "Improved algorithms for movement pattern recognition and classification in physical rehabilitation," in *2019 10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, 2019, pp. 417–424.

[16] G. Sziladi, T. Ujbanyi, and J. Katona, "Cost-effective hand gesture computer control interface," in *2016 7th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, 2016, pp. 000 239–000 244.

[17] V. Szucs, T. Guzsvinecz, and A. Magyar, "Movement Pattern Recognition in Physical Rehabilitation - Cognitive Motivation-based IT Method and T. Guzsvinecz et al. The Cognitive Motivation-based APBMR Algorithm in Physical Rehabilitation," *Acta Polytech. Hung.*, vol. 17, no. 2, pp. 211-235, 2020.

[18] T. Guzsvinecz, V. Szucs, and A. Magyar, "The Cognitive Motivation-based APBMR Algorithm in Physical Rehabilitation," *Acta Polytech. Hung.*, vol. 20, no. 5, pp. 41–60, 2023.

[19] T. Ujbanyi, J. Katona, G. Sziladi, and A. Kovari, "Eye-tracking analysis of computer networks exam question besides different skilled groups," in *2016 7th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, 2016, pp. 000277–000282.

[20] J. Katona, "Electroencephalogram-Based Brain-Computer Interface for Internet of Robotic Things, Cognitive Infocommunications," *Theory and Applications*, pp. 253–275, 2018.

[21] C. Cristina, "Assessing Visual Attention in Children Using GP3 Eye Tracker," in *Proceedings of the 10th IEEE International Conference on Cognitive Infocommunications*, 2019, pp. 343–348.

[22] A. Sudár and B. Ádám, "Csapó: Interaction Patterns of Spatial Navigation in VR Workspaces," in *Proceedings of the 10th IEEE International Conference on Cognitive Infocommunications*, 2019, pp. 611–614.

[23] Z. Kvasznicza, "Teaching electrical machines in a 3D virtual space," in *2017 8th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, 2017, pp. 000385–000388.

[24] I. Heldal and C. Helgesen, "The Digital HealthLab: Supporting Interdisciplinary Projects in Engineering and in Health Education," *Journal of Applied Technical and Educational Sciences*, vol. 8, no. 4, pp. 4–21, 2018.

[25] F. Erdős, R. Németh, and B. Firuza, "Virtual Teamwork in Gamified 3D Environment," *Infocommunications J., Special Issue: Internet of Digital & Cognitive realities*, pp. 15–20, 2023.

[26] C. Rigóczki and A. Damsa, "Kristóf Györgyi-Ambró: Gamification on the edge of educational sciences and pedagogical methodologies," *Journal of Applied Technical and Educational Sciences*, vol. 7, no. 4, pp. 79–88, 2017.

[27] F. Es-sabery, K. Es-sabery, H. Garmani, J. Qadir, and A. Hair, "Evaluation of different extractors of features at the level of sentiment analysis," *Infocommunications J.*, vol. 14, no. 2, pp. 85–96, 2022.

[28] A. Kovari, "CogInfoCom Supported Education: A review of CogInfoCom based conference papers," in *Proceedings of the 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, 2018, pp. 233–236.

[29] R. Pinter, "Sanja Maravic Cisar: Measuring Team Member Performance in Project Based Learning," *Journal of Applied Technical and Educational Sciences*, vol. 8, no. 4, pp. 22–34, 2018.

[30] Q. Yang, X. Tong, L. Schieb, A. Vaughan, C. Gillespie, J. L. Wiltz, S. C. King, E. Odom, R. Merritt, Y. Hong, and M. G. George, "Vital signs: Recent trends in stroke death rates - United States, 2000-2015," *MMWR. Morbidity and mortality weekly report*, vol. 66, no. 35, p. 933, 2017.

[31] E. J. Benjamin, M. J. Blaha, S. E. Chiuve, M. Cushman, S. R. Das, R. Deo, S. D. de Ferranti, J. Floyd, M. Fornage, C. Gillespie, C. R. Isasi, M. C. Jiménez, L. C. Jordan, S. E. Judd, D. Lackland, J. H. Lichtman, L. Lisabeth, S. Liu, C. T. Longenecker, R. H. Mackey, K. Matsushita, D. Mozaffarian, M. E. Mussolino, K. Nasir, R. W. Neumar, L. Palaniappan, D. K. Pandey, R. R. Thiagarajan, M. J. Reeves, M. Ritchey, C. J. Rodriguez, G. A. Roth, W. D. Rosamond, C. Sasson, A. Towfighi, C. W. Tsao, M. B. Turner, S. S. Virani, J. H. Voeks, J. Z. Willey, J. T. Wilkins, J. H. Wu, H. M. Alger, S. S. Wong, and P. Muntner, "Heart Disease and Stroke Statistics'2017 *Update: A Report from the American Heart Association. Circulation*," vol. 135, no. 10, pp. e146–e603, 2017.

[32] E. Benjamin, J. Michael, S. E. Blaha, and M. Chiuve, "Mary Cushman: Heart disease and stroke statistics 2017 update: a report from the American Heart Association," *American Heart Association. Circulation*, vol. 135, pp. 229–445, 2017.

[33] Q. Yang, X. Tong, L. Schieb, and A. Vaughan, "National Center for Chronic Disease Prevention and Health Promotion Division for Heart Disease and Stroke Prevention, MMWR Morbidity and mortality weekly report," in *Recent trends in stroke death rates – United States*, 2000.

[34] B. Bonnechère, B. Jansen, L. Omelina, and S. Van Sint Jan, "The use of commercial video games in rehabilitation: a systematic review: A systematic review," *Int. J. Rehabil. Res.*, vol. 39, no. 4, pp. 277–290, 2016.

[35] N. Hocine, A. Gouaïch, S. A. Cerri, D. Mottet, and J. Froger, "Isabelle Laffont: Adaptation in serious games for upper-limb rehabilitation: an approach to improve training outcomes," *User Model. User-adapt Interact.*, vol. 25, pp. 65–98, 2015.

[36] A. Ghazarian and S. M. Noorhosseini, "Automatic detection of users' skill levels using high-frequency user interface events," *User Model. User-adapt Interact.*, vol. 20, no. 2, pp. 109–146, 2010.

[37] T. Guzsvinecz, V. Szucs, and C. Sik-Lanyi, "Suitability of the Kinect sensor and Leap Motion Controller-A literature review," *Sensors (Basel)*, vol. 19, no. 5, p. 1072, 2019.

[38] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE transactions on information theory*, vol. 13, no. 1, pp. 21–27, 1967.

[39] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, V olume 1: Statistics. Berkeley, CA: University of California Press, 1967, pp. 281–298.

Tibor Guzsvinecz is an Assistant Professor at the Faculty of Information Technology of the University of Pannonia, Hungary. He obtained his PhD degree in the field of Information Science in 2021. His research areas are data science, gamification, human- computer interaction, spatial ability, and virtual reality. He authored 69 scientific publications which include 19 journal papers, and 41 studies found in various conference proceedings. At the moment, he is an editor of a book project about gamification. He is an IEEE member since 2023.

Judit Szűcs is an Assistant Professor at the Faculty of Information Technology of the University of Pannonia, Hungary. She received her PhD degree in 2021 in Computer Science at the University of Szeged. Her main interests include binary tomography, image processing, and computer vision. She has 30 scientific publications which include 6 journal papers, and 19 studies found in various conference proceedings. She is an IEEE member since 2023.

Veronika Szűcs is an Associate Professor at the Faculty of Information Technology of the University of Pannonia, Hungary. She obtained her Ph.D. degree in the field of Information Science in 2019. Her research areas are human-computer interaction, information systems, virtual environments, IT and information systems' security.

Robert Demeter is an Associate Professor at Department of Automation and Computer Science, Faculty of Electrical Engineering and Computer Science (UTBV), Brasov. He is involved in a number of international interdisciplinary research projects in which he contributes to the IT side. His main research interests are ICT-supported education, classification of big data, planning and implementation of communication applications.

József Katona is Associate Professor at University of Dunaujvaros, Obuda University and John von Neumann University, Hungary. He holds a Ph.D. degree in Computer Engineering. His research areas are human-computer interfaces, brain-computer-interfaces, programming, software development, eye-tracking analysis.

Attila Kovari is Professor at Eszterhazy Karoly Catholic University, University of Dunaujvaros, Obuda University and John von Neumann University, Hungary. He holds a Ph.D. degree in Engineering and Habilitation in Computer Engineering and Education. His research areas are human-computer interaction, brain-computer-interfaces, eye-tracking analysis, IT in industrial applications.

Concepts of Cognitive Infocommunications

Péter Baranyi[1,2,3], Borbála Berki[1,2,3,4,*], and Ádám B. Csapó[1,2,3,4,*]

***Abstract*—Cognitive Infocommunications (CogInfoCom) is an interdisciplinary field that explores the interplay between the cognitive sciences and infocommunication technologies (ICT). This paper presents recent transformative changes in Cognitive Infocommunications, emerging both on the technological and the human side. Concepts such as Digital Reality and Cognitive Entities are also presented, which have emerged as a result recent technological convergence and further entanglement of ICT with human cognition. In addition, the current focal points of CogInfoCom research are presented to emphasize the entangled nature of human-technology interactions and human-technology co-evolution.**

***Index Terms*—CogInfoCom, cognitive infocommunications, cognitive entity, digital reality, Generation of Cognitive Entities (Gen CE).**

I. Introduction

In the current era of rapid technological advancements and increasing reliance on digital systems, the field of cognitive infocommunications (CogInfoCom) [1], [2] has emerged as a key area of study. CogInfoCom is an interdisciplinary field that explores the interplay of cognitive sciences and information and communication technologies (ICT), aiming to understand and enhance both short-term interactions and long-term co-evolution between humans and digital systems. In order to support researchers and practitioners in orienting this landscape, in this paper, we take stock of the fundamental concepts behind the fields and organize them into a novel system by highlighting emergent relationships between them. This is important both in terms of establishing a common vocabulary that can prevent the fragmentation of the field and also in terms of increasing the comprehensibility of these key concepts for those who are new to the field or have experience in other, closely related fields such as AI.

Over the past decade, digital technologies, infocommunication devices, as well as the psychological, cognitive and social context in which users encounter them have undergone several changes and developments. The co-evolution of infocommunication technologies and humans seems to have reached the point where a unified perspective focusing on cognitive entities as a whole seems to be a more viable research approach in many cases compared to existing approaches focusing on separate human and techical measures.

[1] Corvinus Institute for Advanced Studies, Corvinus University of Budapest, Hungary
[2] Institute of Data Analytics and Information Systems, Corvinus University of Budapest, Budapest, Hungary
[3] Hungarian Research Network, HUN-REN, Budapest, Hungary
[4] Óbuda University, Budapest, Hungary
* E-mail: {borbala.berki, adambalazs.csapo}@uni-corvinus.hu

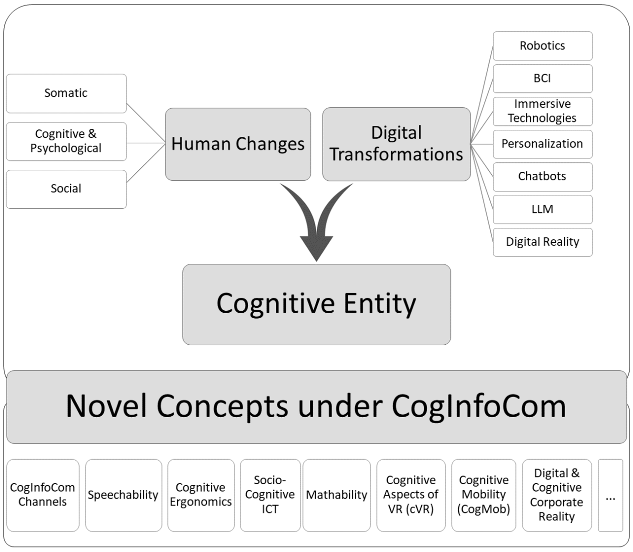


Fig. 1. A graphical overview of the structure of the paper and the key concepts discussed in the paper.

The paper is structured as follows. Section II presents some of the recent changes that we have identified in human users as a result of the widespread adoption of ICT technologies. This is followed by a presentation of recent changes in technology which are relevant to CogInfoCom in Section III. We make the case that these changes are so entangled, not just on the technological level, but also involving the human users that the following part it is worth highlighting the concept of Cognitive Entities, which is the subject of Section IV. This novel focus on the notion of cognitive entities in turn underlines the need to revisit other key concepts of CogInfoCom; therefore, Section V presents some of the more recent focal points of CogInfoCom in light of these developments. The paper concludes with a brief summary of its key points in Section VI. The structure of the paper is presented visually on Fig. 1.

II. Changes in Human Physical and Cognitive Systems

In recent decades, human behavior and lifestyles have undergone significant changes due to the widespread adoption of various infocommunication technologies. These changes have brought about irreversible transformations, particularly concerning the accessibility of the Internet, even from early childhood and in increasingly remote locations, leading to distinct developmental patterns compared to previous generations.

Although individual usage patterns, ergonomic practices, and personal habits vary, there has been a general increase

DOI: 10.36244/ICJ.2024.5.5

in average daily screen time. Since 2013, this has risen by nearly fifty minutes. As of 2021, the average daily screen time across all devices is 6 hours and 58 minutes [3]. The pervasive use of mobile devices for activities such as voice/e-mail communication, social media, multimedia consumption, and gaming has significantly impacted body posture; an effect that is noticeable both in terms of ergonomics and over the long term. Furthermore, these novel usage patterns of ICT devices also have broader implications for users' mental and physical health.

Below, we summarize some of the somatic, cognitive / psychological and sociel changes that are most relevant to the topic at hand.

A. Somatic effects

1) Tech Neck: Pain symptoms in the neck and neighboring regions are commonly experienced during or after prolonged looking down, such as when engaging in text messaging on handheld devices [4]–[6]. Research suggests that the use of smartphones while standing results in a more rigid posture compared to sitting [7]. These neck and shoulder pain symptoms are not limited to adults; they also affect teenagers [8]. In addition to phone usage, other factors such as prolonged sitting, obesity, and sedentary lifestyle have been identified as contributors to the increased prevalence of musculoskeletal pain [5].

2) Texting Thumb: The so-called "texting thumb", "PlayStation thumb", or more precisely, De Quervain's syndrome, is a common repetitive use injury that is frequently observed. Studies have found a higher prevalence of this syndrome among individuals with problematic smartphone use [9].

3) Vision Problems: Excessive screen time and prolonged staring at screens can lead to various vision problems, including eye strain, dryness, blurred vision, and headaches. These symptoms are often associated with conditions such as computer vision syndrome and digital eye strain. Additionally, the blue light emitted by screens can disrupt sleep patterns and potentially contribute to long-term eye problems [5], [10].

4) Sleep Disruptions: Previous research on screen time [11] has already emphasized the importance of minimizing screen usage, especially in the hours leading up to bedtime. However, the widespread use of smartphones has made it common for individuals to use them directly before sleep, which can further disrupt sleep patterns. One contributing factor is exposure to blue light, which can suppress the production of melatonin, a hormone that regulates sleep-wake cycles. The constant connectivity and notifications from smartphones can also disrupt sleep by inducing anxiety and promoting a state of alertness rather than relaxation. In a recent experiment, it was found that reducing blue light exposure during the night resulted in improved subjective sleep quality [12]. Additionally, other results indicated that using smartphones in bed led to increased sleep latency, higher average heart rate, lower heart rate variability, and a decrease in total sleep time. These findings suggest a poorer quality of rest, with a higher proportion of wake periods during the night [13].

B. Cognitive and Psychological Changes

The constant use of technology has been found to have both positive and negative effects on brain function and behavior [14]. The key contributing factors are information overload, screen time, and media multitasking. Interestingly, information overload has been a long-standing issue throughout history, but in the 21st century, its consequences are more pronounced due to society's heavy reliance on information [15]. Media multitasking, characterized by simultaneous processing or, according to the neurocognitive literature, rapid task-switching, has become a prominent feature of modern technology [16]. In the following, we consider different aspects of the impacts these changes have on human cognition.

1) Cognitive Development: Emerging research on the impact of technology on children reveals intriguing and slightly contradictory findings.

In general, children's screen time has become a concern, as even children under the age of 2 are now spending over an hour every day engaged with screens, leading to reduced reading time. This pattern has been associated with negative outcomes such as impaired language development, executive functioning, and decreased connectivity between brain regions involved in word recognition, language, and cognitive control. Additionally, screen time has been found to predict behavioral problems and poorer theory of mind [14], [17].

At the same time, studies have also uncovered several factors that can mitigate or sometimes reverse these tendencies. For example, studies have suggested that high-quality television programs without commercials can even be positively associated with attention and executive control measures. Conversely, commercial breaks disrupt children's engagement, indicating the importance of uninterrupted content [18], [19]. Educational cartoons and programs have also been linked to positive learning outcomes and improved executive function [20]. Moreover, touch screen devices, with their interactive nature, facilitate active engagement and control over information flow, enhancing children's ability to learn from video content [21].

These insights underscore the significance of content quality and interactivity in shaping the effects of technology on children's attention, learning, and cognitive development [16].

Furthermore, technology can offer cognitive exercise for the aging brain. Engaging in mentally challenging tasks such as searching online has been shown to potentially delay cognitive decline [22]; as a result, serious gaming can play a key role in mitigating the effects of aging, both physical and cognitive [23]–[25].

2) Attention: Reduced attention span, originally noted by Carr [26] on the basis of self-observation, is a widely discussed topic, as some users have perceived a decline in their attention, both in terms of span and intensity. Although the direct link with the technology remains to be conclusively demonstrated, experts generally agree that a causality exists. The constant flow of information could interfere with sustained concentration and promote multi-tasking between various information sources [27]. It is also possible that repetitive attentional shifts and multitasking directly impair human executive functions [28], and can also be linked to changes in the locality versus globality of attention [17]. Another factor of influence on attention span relates to the limited opportunities for offline interaction due to the constant use of technological devices. This lack of offline engagement may hinder the brain's ability to rest in its default mode, leading to attentional problems [29].

In addition to the above, structural changes have observed in individuals with Internet Use Disorders in the brain regions associated with attentional control, compared to the healthy control group [30]. These individuals have a threefold increased likelihood of Attention Deficit Hyperactivity Disorder (ADHD) [31]. Reduced gray matter in the anterior cingulate cortex and other prefrontal regions that are related to sustained concentration and the ability to ignore distractor stimuli has been observed in connection with excessive Internet usage [32] and media multitasking [33].

On the other hand, improved visual and spatial attention has also been observed in users after playing action video games for more than 4 days per week for 6 months [34]. In a similar vein, surgeons who played video games made 37% fewer surgical errors, and their response time was 27% faster than surgeons who did not play video games in one study [35].

3) Memory & Information Retrieval: There is a growing concern regarding the overreliance on online information as it serves as an always available external memory storage [31], [36]. However, certain patterns of brain activation that are crucial for the long-term storage of registered information are found to be missing [37], [38]. This raises questions about the potential impact of these phenomena on memory and cognition. One concept that has been explored is the "Google effect", where individuals tend to remember less factual information but instead remember where to access that information [37]. This suggests a shift in cognitive processes as individuals rely on the availability and accessibility of online information rather than their own memory.

4) Digital Skills: A study with a sample of 191 school-aged children categorized as either high-digital users or low-digital users, assessed these children in terms of verbal and visuoperceptual cognitive performance through standardized tests and a self-report questionnaire [39]. The results demonstrated a positive impact of digital exposure on cognitive development, as the high-digital users outperformed low-digital users in naming, semantic, visual memory, and logical reasoning tasks. These findings align with existing literature, highlighting the significant role of technology in enhancing cognitive abilities and promoting smart learning among children.

5) Multitasking Skills: Despite the negative impact multitasking can have on human cognition, it has become a skill valued in today's fast-paced world. However, it is important to note that multitasking abilities tend to decline linearly across the lifespan. Interestingly, experimental research has indicated that certain computer games have the potential to enhance multitasking skills [14], [40], [41].

6) Emotional & Social Intelligence: Increased screen time is often associated with reduced face-to-face communication [14]. Furthermore, a study conducted on preteens demonstrated that after five days without screens, they exhibited improved recognition of nonverbal emotional and social cues [42]. Another review by Bochicchio et al. [43] revealed the complexity of the role of digital games in children's cognitive and socio-emotional development, including both positive effects such as pro-social behaviors and negative effects such as anti-social behaviors and isolation. These findings highlight the potential impact of technology addiction on social interaction and the importance of exploring mental health apps as potential interventions [14].

C. Social Changes

Besides the technological and cognitive changes outlined in previous sections, it is also important to consider the social context in which users spend their everyday life. Many of the previously mentioned changes imply a change on the societal level. Connectivity without geographical limits among people has increased with the widespread access to smartphones and the Internet. Instant communication creates the feeling of connectedness. Digital communities have arisen as a result of these new means of communication and interaction.

The overall effects of these socio-technological changes are greatly under-researched at this time. As our social interactions intertwined with cognitive and technological changes reach an increasingly high level of complexity, not much data is available that could help uncover causal phenomena, leaving room for most researchers and practitioners to merely speculate. However, building on Heidegger, philosophers have raised the question as to whether technology is or can be regarded as neutral. For instance, the South-Korean philosopher Byung-Chul Han argues that the socia-technical situation today is responsible for burnout, for a significant increase in narcissism, for the disappearance of rituals, and for a general alienation. He argues that our societies followed not the Orwellian but rather the Huxleyan path towards a negative utopia, such that we have created a panopticon in which we are our own prison guards [44].

III. Digital Transformations

In addition to the human changes described in the previous section, digital technologies have undergone significant advancements in recent years. The digital transformations of the past decade have enhanced the cognitive capabilities of information and infocommunication systems. These systems, often referred to as "smart", have become increasingly complex, exhibiting advanced functionalities.

Furthermore, with the rapid pace of technological innovation, the issue of technology adoption has become increasingly relevant. The initial phase of technological acceptance is characterized by high expectations placed on new technologies, which may initially be exaggerated but eventually settle into a more realistic level. Criticisms during this phase play a vital role in guiding future developments by addressing shortcomings and shaping the trajectory of improvement. These evaluations serve as valuable feedback that help refine and optimize emerging technologies toward their full potential.

A. Robotics

Recent advances in robotics have been transformative, expanding the capabilities of robots. Notably, autonomous navigation in complex and dynamic environments has seen remarkable progress, thanks to the integration of sensor technologies and AI algorithms [45]. Learning-based methods are employed for ground vehicles navigating unstructured environments, utilizing perception to achieve context-aware navigation [46]. The emergence of collaborative robots, or cobots, has facilitated physical interaction between humans and robots in shared environments, opening up new possibilities in manufacturing, logistics, and healthcare [47]. Surgical and rehabilitation robots have made a significant contribution to the field of medicine, enhancing precision and patient outcomes [48]. Additionally, the development of soft robotic systems, that can interact with delicate objects more effectively can lead to further clinical applications [49]. By leveraging perception, navigation, and cognitive science [50], robotics continues to advance, pushing the boundaries of what robots can achieve in various domains.

B. Brain-computer Interfaces

Brain-computer interfaces (BCIs) facilitate direct communication between the brain and external devices, benefiting individuals with disabilities and there are further applications in diverse fields. BCIs receive and interpret brain signals, (and in the case of paralyzed patients it bypasses impaired neuromuscular pathways), to accomplish tasks and interact with the environment. Thanks to rapid advancements in neurotechnology and AI, brain signals used for brain-computer communication have progressed beyond sensory levels to encompass higher-level cognitive functions like goal-directed intentions [51], [52]. Advancements in flexible electronics have made BCIs compatible with the brain's mechanical properties [53]. The I3 evolutionary model [51] for BCIs includes stages of interface, interaction, and intelligence. Brain-computer interaction enables bidirectional communication and co-adaptation between the brain and computer. The integration of AI technology could augment human intelligence with BCIs [51].

C. Immersive and Non-Immserive Spatial Technologies

Over the past decade, Virtual Reality (VR), Augmented Reality (AR), and Extended Reality (XR) have witnessed remarkable advancements, revolutionizing the way we perceive and interact with digital content. VR and AR devices improved significantly in terms of display resolution, tracking accuracy, ergonomics, and wearable sensors [54], [55]. These improvements led to the application of these technologies in various industries such as healthcare [56], [57], education [58], [59], architecture [60], tourism [61] and entertainment. In an industrial setting, digital twins coexist with physical entities: the virtual model is connected to the physical one by sensors, to generate real-time data [62]. Digital twins allow the implementation of virtual safety training too [63]. Another current topic in this field is collaboration [64] and user-friendliness of content creation tools, that could allow the users to more easily create or edit the virtual environments [65], [66]. Similarly, the personalization of virtual spaces has become a topic of interest [67].

D. Personalization

Personalized recommendation systems have gained widespread usage across various domains, aiming to enhance user experiences by offering relevant and customized suggestions. These systems leverage user preferences and behavior to provide tailored recommendations. By analyzing user interactions with the system, such as past choices, browsing history, and feedback, these recommendation systems offer suggestions that align with individual interests [68]. As the demand for increasingly personalized experiences grows, certain systems go beyond conventional approaches, towards hyperpersonalization which involves real-time matching to provide customized experiences [69]. However, it should be noted, that hyperpersonalization can also trigger negative reactions [70]. These systems have the power to influence which information is easily accessible to users, effectively serving as digital nudges. Personalized recommendations impact our decision-making processes, as they guide and shape our choices based on the information presented [71]. This is often seen as beneficial, however, it has also been mentioned by several authors that optimizing a cost function in such a way can also have unexpected effects: by modifying the data (e.g., changing and leading to a uniformization of human behaviors) rather than predicting the data (e.g., understanding human preferences) more precisely [72]. As a result, Dörfler [73] emphasizes that although pattern recognition and its use in personalization can be hugely effective from a business standpoint, in a broader sense it does not work nearly as well as it is showcased by vendors: offten, instead of offering people what they want, they tell them what they *should* want.

E. Chatbots

Interest in chatbots has grown rapidly since 2016, with advancements in NLP allowing them to provide more interactive experiences [74]. Chatbots and virtual assistants are computer programs that simulate human-like conversations and understand human languages. They have found applications in diverse fields such as marketing, education, healthcare, cultural heritage, and entertainment. By combining NLP with cognitive computing, chatbots can offer intelligent functionalities across different categories, contextualizing data, personalizing information, and abstracting it e.g. for decision-making [75]. Acceptability and effectiveness of chatbots varies depending on the context, with factors like empathy in healthcare and pleasure in entertainment playing important roles [76]. Overall, chatbots with NLP capabilities enhance interactions and provide tailored experiences in various domains.

F. Large Language Models

Recent breakthroughs in AI research have been driven by the development of large language models (LLMs) based on the Transformer architecture [77]. Trained on massive amounts of web-text data, LLMs excel in natural language processing tasks. GPT-4 represents progress toward artificial general intelligence (AGI), as it can seemingly understand

and connect any topic, and perform diverse tasks. However, it is important to note that current capabilities still fall short of human-level abilities [78], especially when it comes to reasoning across multiple domains.

In the recent past, a new approach to studying LLMs has emerged, drawing inspiration from and being based on traditional psychology rather than solely relying on machine learning techniques [78]. An example of such an approach is the investigation of Theory of Mind (ToM) in LLMs [79]–[81], which refers to the ability to understand and attribute mental states to others. Current results showed that current LLMs seem to rely on shallow heuristics rather than robust ToM abilities [80]. A similar approach was applied with the creation of the "Baby Intuitions Benchmark (BIB)" tasks which probe commonsense psychology in LLMs. Results revealed that current neural-network models failed to capture infants' knowledge in many ways [82].

G. Digital Reality

Besides the novel advances in different technologies, an equally important aspect is their entanglement and convergence. Various information technologies, once seen as separate domains, are now becoming increasingly interconnected. Virtual Reality, Augmented Reality, Mixed Reality, Digital Twins, Artificial Intelligence, 5G networks, and the omnipresent 2D Web are experiencing widespread adoption, and their integration has the potential to shape a fundamentally new reality. This integration blurs the boundaries between the physical and digital world, as well as digital representations and simulations, leading to the emergence of a concept known as Digital Reality [83], [84]. This is a "*high-level integration of virtual reality (including augmented reality, virtual and digital simulations and twins), artificial intelligence and 2D digital environments which creates a highly contextual reality for humans in which previously disparate realms of human experience are brought together*". [83].

From the user's perspective, the convergence of these technologies feels seamless, as they complement one another and give rise to a digital reality. The boundaries between different technologies are becoming increasingly blurred, enabling easier connection and alignment, thereby facilitating more intuitive usage. The advent of Language and Learning Models (LLMs) further enhances this experience, as language serves as a universal interface, akin to a meta-dimension (metaD), enabling natural communication for human users.

IV. Cognitive Entity

Based on previously presented changes and developments both on the human and technological side, it is not surprising that the relationship between humans and technology has also become more engangled in recent years [2], [85]. As both human capabilities and technological advancements continue to evolve, their collaboration has the potential to yield a new level of entangled cognition, characterized by a qualitatively different experience. The concept of cognitive entities was introduced as a "*synergetic combination of humans, devices, infrastructure, and environment that is identifiable from the perspective of some (high-level) cognitive capability*" [2]. This term captures the entangled sets of humans and information and communication technology (ICT) capabilities that have evolved together [86]. While humans can be cognitive entities in themselves, many theoreticians and philosophers view machines as being incapable of cognitition without humans. For example, John Searle makes the distinction between *observer-independent* and *observer-relative* intelligence such that, for example, one chess grandmaster overcoming another chess grandmaster is a display of intelligence in an observer-independent sense, whereas Deep Blue overcoming Kasparov is a display of intelligence in an observer-relative sense: it is the human user of the computer who interprets its outputs as moves while the computer does not even "know" that it is playing chess [87], [88]. Building on the work of Hubert Dreyfus and John Searle, Dörfler [73] explores in which areas AI can replicate human performance or even outperform humans and what the areas are that are uniquely human.

The collaboration between humans and infocommunication technology facilitates joint operations that lead to increased efficiency. One illustrative example of this collaboration is when mathematicians work alongside current programs, resulting in a synergistic brainstorming process. In such instances, if the machine does not provide a counterexample, it signifies the potential value of exploring the idea further. Conversely, if the machine quickly identifies an existing counterexample, it indicates that pursuing the idea may not be worthwhile. This collaborative approach harnesses the strengths of both human intuition and machine processing power.

The concept of "generation CE" (generation of cognitive entities) was introduced in 2015 in [2] and later extended in [86]. Since then, the entanglement between humans and ICT has grown tighter and become inseparable as presented in the previous section. New, digital skills have emerged, such as learning and communicating through generative AI. Also in cognitive tasks such as naming, semantic and visual memory, and logical reasoning, high-digital users have been shown to outperform low-digital users [39]. This generation has grown up in close relation with information and communication technology from early childhood, leading to a genuine co-evolution with technology. Unlike older generations that had to adapt to emerging technologies, members of Generation CE have a more inherent relationship with ICT due to their early exposure.

Generation CE includes those born after 2010. This generation experiences a seamless integration of cognitive ICT into their everyday lives, with extended cognitive capabilities and easy access to information through devices at their fingertips. Their social lives are also partially conducted online, with online communities holding comparable significance to real-life communities.

A. Levels of Entanglement

From the perspective of interaction modes, the convergence of humans and information and communication technology can be observed across three distinct levels [86].

The first level of entanglement pertains to direct relationships at a low level, involving invasive and non-invasive interfaces such as brain-computer interfaces. This level enables direct sensing and control, however in most cases it is cumbersome due to the need for implanted or worn sensors, making it challenging to operate. Nevertheless, current technological advances allow us to go beyond sensory levels to encompass higher-level cognitive functions (as presented in Section III-B). The I3 model shows further directions in the direct linking of artificial and natural intelligence. Augmenting human intelligence by using BCI technology would allow the improvement of human perceptual abilities, information processing, and decision-making. The co-adaptation between brain and computer is a fundamental issue in the development of BCIs [51].

At the level of personal informatics devices, a different form of entanglement arises, where communication and interaction occur through various sensory modalities, including human and non-human components. Determining the appropriate "communication language" for message encoding, considering the semantics of the information, modalities used, application environment, and users' cognitive capabilities, becomes crucial at this level. Accommodating the transfer of a wide range of semantic concepts within the limitations of human sensory modalities presents a significant challenge. The trends in personalizing and providing relevant and customized suggestions allow an even higher reliance on these devices.

Lastly, a third level of entanglement manifests at the collective level of multi-user interactions. Applications at this layer leverage collective behaviors to support individual user interactions or analyze past behaviors (both individual and collective) to predict or analyze collective events. These applications often involve data-mining and analyzing diverse data sources, including social communication platforms and various activities.

B. Joint Measurement

If we take a more distant perspective, the close cooperation between humans and information and communication technology (ICT) allows for the possibility of joint measurement of the human-ICT combination. As humans and ICT co-evolve, new methodological considerations have emerged due to their interconnectedness becoming an integral aspect of our daily lives. Traditionally, human cognitive abilities and artificial cognitive abilities were assessed separately using diverse methods, although in some cases, the methods have converged, as presented in Subsection III-F. Given the potential for joint operations between humans and ICT to enhance efficiency, it is pertinent to measure the combined cognitive capability of this human-ICT combination rather than focusing solely on either component. The emphasis should shift toward measuring the cognitive entity as a whole. The Google effect [37] described earlier also shows the need for this shift in perspective, as it is so natural that the Internet is available at any time that people remember where they can access information rather than the factual information itself. Measuring cognitive entities instead of only natural cognitive capabilities would make it possible to measure general, everyday cognitive abilities since people are now almost only separated from their various smart devices in laboratory conditions.

C. Balanced Entanglement

The advantages of close entanglement and collaborative operation are easy to understand, as this allows a more efficient way to resolve related tasks. However, in some cases, there could be drawbacks to this close relationship. For example, turn-by-turn navigational aids that are commonly found in GPS devices have a negative effect on human spatial abilities. This detrimental effect can be lowered, by raising awareness and being more mindful when deciding whether to rely on navigational aids or not. Additionally, optimizing the interaction design of these devices to facilitate active information processing and integrating landmarks as reference points in GPS aids can promote greater environmental awareness [89]. Thus, it is important to find the right balance and to use these advantages of interconnectedness wisely.

V. Novel Concepts under CogInfoCom

Given the significant changes occurring on multiple levels, it becomes necessary to reassess the concept of CogInfoCom and its subtracks, placing greater emphasis on the role of cognitive entities. Cognitive infocommunications (CogInfoCom) [1], [2] investigates the "*link between the research areas of infocommunications and cognitive sciences, as well as the various engineering applications which have emerged as a synergic combination of these sciences*". Its primary focus is on the synergistic integration of these disciplines and their practical applications. The overarching goal of CogInfoCom is to explore the co-evolution of cognitive processes with infocommunications devices, enabling the extension of human cognitive capabilities beyond geographical constraints and fostering collaboration between natural and artificial cognitive systems. By facilitating this integration and expansion, CogInfoCom aims to enhance the effectiveness of engineering applications that involve cognitive collaboration.

It is important to emphasize that all presented concepts are intricately intertwined and entangled, influencing each other. They do not exist as isolated entities; rather, they coexist while maintaining their individual significance and impact. The current focal points of CogInfoCom, are presented in Fig. 2. The outer circle illustrates the three recent novel directions, which have seen significant advancements in the last few years. This progression is attributed to a substantial technological breakthrough, enabling their quicker integration into applied environments. Notably, these directions offer a more tangible manifestation and, in these cases, a rapid convergence with business and industrial demands aligning seamlessly with the ongoing research trajectory.

A. CogInfoCom Channels

CogInfoCom channels [90] tackle the communication challenges among cognitive entities when conveying new information. The framework integrates structural and semantic elements to define sensory messages linked to high-level concepts, utilizing icon-like and message-like design elements across modalities [91]–[96]. It incorporates a concept algebra-based toolset and the spiral discovery method [97] for mapping semantic meaning.

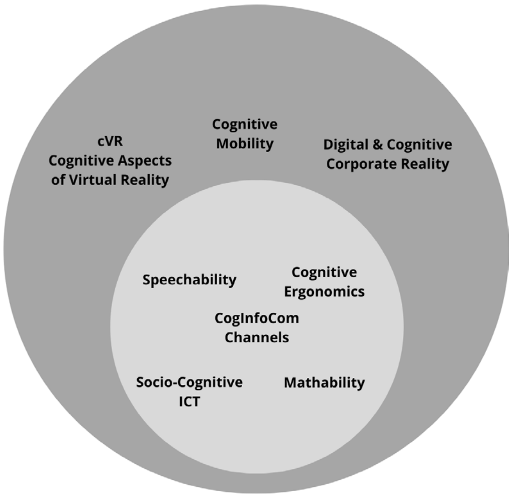


Fig. 2. Graphical representation of the current focus point under CogInfoCom. Note that the representation is not meant to be exhaustive; in addition to these concepts there exist others within the field of CogInfoCom that are still relevant and worth exploring.

B. Speechability

Speech and language are a fundamental part of human interaction. Speech is a complex modality, it has verbal components and also other components such as prosody, gestures, facial expressions, and so on. Furthermore, in human communication, speech not only serves for information sharing but also the emotional foundations and building of trust. The term *speechability* refers to the perspective of artificial devices, which are capable of generating speech (as presented in Subsection III-F about LLMs), but it is not the way humans are capable of doing it [98]. Several CogInfoCom researchers, including Anna Esposito, Carl Vogel, Gennaro Cordasco, Maria Koutsombogera, Stanislav Ondas, Klára Vicsi, and Costanza Navarretta, focused on the different aspects of speechability, such as speech recognition, multimodal cues, conversational fillers, phonological and spectral analysis and also corpus-based techniques [99]–[111].

C. Cognitive Ergonomics

Cognitive ergonomics, a field closely related to Human-Computer Interaction (HCI), encompasses the study of multi-sensory usability challenges arising from the increasing use of augmentative reality mediation. Noteworthy research has been conducted by Karoly Hercegfi, Anita Komlodi, and their research groups, exploring the opportunities and requirements in the field of ergonomics and human factors. Their work has highlighted the adaptation of ergonomic approaches, leading to synergies with CogInfoCom [112]–[116]. Furthermore, Ágoston Török [117] discussed the shift from HCI towards cognitive infocommunications, emphasizing understanding human behavior, limits, needs, and cognition in interface design. Additionally, Molnár and colleagues [118], [119] investigated the connection and communication between entities in infocommunication. Many researchers have also contributed to the diverse fields of user interface design, several researchers have also contributed, such as Atsushi Ito, Mihoko Niitsuma, Miroslav Macik and Zdenek Mikovec [120]–[123].

D. Socio-Cognitive ICT

Socio-Cognitive ICT, proposed by Hassan Charaf and his research group, refers to the intersection of social and cognitive aspects within computer networks. It emphasizes the cognitive properties that emerge through user interactions and explores how analyzing and managing information flow can enhance social capabilities. This concept enables applications like understanding crowd-generated phenomena and optimizing information flow in critical situations, bridging the gap between the social and cognitive aspects of ICT.

The term “socio-cognitive ICT” has gained prominence in describing these applications, which leverage the ongoing evolution towards the Internet of Digital & Cognitive Realities (IoD) and Smart Ecosystems [124], [125]. As Sallai [126] underlined, the convergence of communications [127], [128], information technology, media, Internet of Things (IoT), AR, VR, and artificial intelligence plays a significant role in creating smarter environments, including Smart Cities and Smart Factories.

E. Mathability

Mathability is a research direction that explores artificial and natural cognitive capabilities relevant to mathematics. It was introduced by Baranyi and Gilányi in 2013 [129], and through the years it evolved and broadened its meaning thanks to the work of Katarzyna Chmielewska [130]. It encompasses various aspects, ranging from basic arithmetic operations to high-level symbolic reasoning [129]. In addition to computing quantitative answers, mathability also emphasizes the ability to reason about the processes behind reaching those answers and incorporates estimated quantities and simplifying assumptions. This interdisciplinary field draws foundations from both human-related cognitive sciences and technology [131]–[136].

F. cVR – Cognitive Aspects of Virtual Reality

Cognitive Aspects of Virtual Reality (cVR) [137] is a field that explores the integration of 3D virtual environments, human spatial cognition, AI, and the digital world to enhance and augment human capabilities [138]–[144], particularly in understanding geometric, temporal, and semantic relationships, with implications across sectors such as education [145]–[154], commerce [155], [156], healthcare [57], [157], [158], and industrial production [159]–[161].

Recent advancements in virtual reality have shown promising applications in psychotherapy, including increasing empathy in individuals with Personality Disorders through shared emotional experiences [162] and improving the treatment of Social Anxiety Disorder through cost-effective and efficient virtual reality exposure therapy [163], highlighting the expanding potential of cVR in mental health.

cVR is a great example of how a branch of CogInfoCom is starting to grow and become a field in its own right: there is now a conference on the topic and more and more research institutes and universities are exploring the field.

G. Cognitive Mobility – CogMob

Cognitive Mobility (CogMob) integrates mobility, transportation, AI, and social sciences, aiming to understand and optimize mobility as a blend of artificial and human cognitive systems [164], [165]. Its primary objective is to present a comprehensive perspective on comprehending, modeling, and enhancing mobility on a broader scale, which provides a new holistic viewpoint of mobility. Notably, CogMob places emphasis on engineering applications within the mobility domain, aligning with its inherent nature.

H. Digital & Cognitive Corporate Reality

Digital & Cognitive Corporate Reality (DCR) is an interdisciplinary field that combines corporate management, business science, Internet of Digital & Cognitive Reality (IoD), and cognitive infocommunications to achieve a holistic understanding [166]. It explores the interactions among digital corporate ecosystems, IoD approaches, and hybrid human, organizational, and artificial cognitive capabilities. It aims to develop theoretical frameworks and practical solutions for various applications in this domain [167]–[171].

VI. Conclusion

The field of cognitive infocommunications has undergone significant changes in the last decade. Both human capabilities and digital/infocommunication technologies have seen remarkable transformations and advancements. As a result, not only has human psychology, cognition and the human social context changed, but a novel concept of Cognitive Entities has also emerged, in which human and digital capabilities are no longer clearly separable. The convergence of technologies has given rise to the concepts of Digital Reality, Cognitive Aspects of Virtual Reality and Cognitive Mobility. These developments call for a reevaluation of the field of cognitive infocommunications, and have motivated our presentation of its current focal points and branches.

Acknowledgment

The authors would like to express their gratitude to the entire CogInfoCom community for their invaluable contributions, support, and unwavering dedication to advancing the field of cognitive infocommunications over the years.

References

[1] P. Baranyi and A. Csapo, “Definition and synergies of cognitive infocommunications,” *Acta Polytechnica Hungarica*, vol. 9, no. 1, pp. 67–83, 2012.

[2] P. Baranyi, A. Csapo, and G. Sallai, *Cognitive Infocommunications (CogInfoCom)*. Springer International Publishing, 2015.

[3] S. Zauderer, “Cross river therapy – screen time statistics,” 2023, accessed on June 12, 2023, https://www.crossrivertherapy.com/research/screen-time-statistics.

[4] S. Khattak, M. Gul, H. Ali Kakar, G. Ullah, and M. Ur Rahman, “The cost of long-term use of smart phones in the form of text neck syndrome; a systematic review,” *Rehman Journal of Health Sciences*, vol. 1, no. 1, pp. 3–5, 2019.

[5] A. Wade, “The impact of digital media use on the musculoskeletal and visual systems,” in *Technology and Adolescent Health*. Elsevier, 2020, pp. 365–380.

[6] S. Neupane, U. T. I. Ali, and A. Mathew, “Text neck syndrome – systematic review,” *Imperial Journal of Interdisciplinary Research*, vol. 3, no. 7, pp. 141–148, 2017.

[7] C. D'Anna, M. Schmid, and S. Conforto, “Linking head and neck pos- ture with muscular activity and perceived discomfort during prolonged smartphone texting,” *International Journal of Industrial Ergonomics*, vol. 83, p. 103134, may 2021.

[8] B. Wirth, T. Potthoff, S. Rosser, B. K. Humphreys, and E. D. de Bruin, “Physical risk factors for adolescent neck and mid back pain: a systematic review,” *Chiropractic & Manual Therapies*, vol. 26, no. 1, sep 2018.

[9] V. A. Benites-Zapata, V. E. Jiménez-Torres, and M. P. Ayala-Roldán, “Problematic smartphone use is associated with de quervain's tenosynovitis symptomatology among young adults,” *Musculoskeletal Science and Practice*, vol. 53, p. 102356, jun 2021.

[10] K. Kaur, B. Gurnani, S. Nayak, N. Deori, S. Kaur, J. Jethani, D. Singh, S. Agarkar, J. R. Hussaindeen, J. Sukhija, and D. Mishra, “Digital eye strain – a comprehensive review,” *Ophthalmology and Therapy*, vol. 11, no. 5, pp. 1655–1680, jul 2022.

[11] L. Hale and S. Guan, “Screen time and sleep among school-aged children and adolescents: A systematic literature review,” *Sleep Medicine Reviews*, vol. 21, pp. 50–58, jun 2015.

[12] P. Randjelović, N. Stojanović, I. Ilić, and D. Vučković, “The effect of reducing blue light from smartphone screen on subjective quality of sleep among students,” *Chronobiology International*, vol. 40, no. 3, pp. 335–342, feb 2023.

[13] S. Kheirinejad, A. Visuri, D. Ferreira, and S. Hosio, ““leave your smartphone out of bed”: quantitative analysis of smartphone use effect on sleep quality,” *Personal and Ubiquitous Computing*, vol. 27, no. 2, pp. 447–466, nov 2022.

[14] G. W. Small, J. Lee, A. Kaufman, J. Jalil, P. Siddarth, H. Gaddipati, T. D. Moody, and S. Y. Bookheimer, “Brain health consequences of digital technology use,” *Dialogues in Clinical Neuroscience*, vol. 22, no. 2, pp. 179–187, jun 2020.

[15] D. Bawden and L. Robinson, “Information overload: An introduction,” jun 2020.

[16] M. Vedechkina and F. Borgonovi, “A review of evidence on the role of digital technology in shaping attention and cognitive control in children,” *Frontiers in Psychology*, vol. 12, feb 2021.

[17] V. Konok, K. Liszkai-Peres, N. Bunford, B. Ferdinandy, Z. Jurányi, D. J. Ujfalussy, Z. Réti, Á. Pogány, G. Kampis, and Á. Miklósi, “Mobile use induces local attentional precedence and is associated with limited socio-cognitive skills in preschoolers,” *Computers in Human Behavior*, vol. 120, p. 106758, 2021.

[18] T. M. Hudon, C. T. Fennell, and M. Hoftyzer, “Quality not quantity of television viewing is associated with bilingual toddlers’ vocabulary scores,” *Infant Behavior and Development*, vol. 36, no. 2, pp. 245–254, apr 2013.

[19] A. I. Nathanson, F. Aladé, M. L. Sharp, E. E. Rasmussen, and K. Christy, “The relation between television exposure and executive function among preschoolers.” *Developmental Psychology*, vol. 50, no. 5, pp. 1497–1506, 2014.

[20] S. M. Fisch, Children's Learning From Educational Television. Routledge, apr 2014.

[21] H. L. Kirkorian, K. Choi, and T. A. Pempek, “Toddlers’ word learning from contingent and noncontingent video on touch screens,” *Child Development*, vol. 87, no. 2, pp. 405–413, mar 2016.

[22] G. W. Small, T. D. Moody, P. Siddarth, and S. Y. Bookheimer, “Your brain on google: Patterns of cerebral activation during internet searching,” *The American Journal of Geriatric Psychiatry*, vol. 17, no. 2, pp. 116–126, feb 2009.

[23] M. Janhunen, N. Katajapuu, J. Paloneva, K. Pamilo, A. Oksanen, H. Keemu, M. Karvonen, M. Luimula, R. Korpelainen, T. Jämsä et al., “Effects of a home-based, exergaming intervention on physical function and pain after total knee replacement in older adults: a randomised controlled trial,” *BMJ Open Sport & Exercise Medicine*, vol. 9, no. 1, p. e001416, 2023.

[24] T. T. H. Nguyen, D. Ishmatova, T. Tapanainen, T. N. Liukkonen, N. Katajapuu, T. Makila, and M. Luimula, “Impact of serious games on health and well-being of elderly: A systematic review,” in *Proceedings of the Annual Hawaii International Conference on System Sciences, ser. HICSS*. Hawaii International Conference on System Sciences, 2017.

[25] J. Li, X. Xu, T. P. Pham, Y.-L. Theng, N. Katajapuu, and M. Luimula, “Exergames designed for older adults: a pilot evaluation on psychosocial well-being,” *Games for health journal*, vol. 6, no. 6, pp. 371–378, 2017.

[26] N. Carr, *The Shallows: What the Internet Is Doing to Our Brains*. WW Norton & Company, 2011.
[27] J. Firth, J. Torous, B. Stubbs, J. A. Firth, G. Z. Steiner, L. Smith, M. Alvarez-Jimenez, J. Gleeson, D. Vancampfort, C. J. Armitage, and J. Sarris, "The "online brain": how the internet may be changing our cognition," *World Psychiatry*, vol. 18, no. 2, pp. 119–129, may 2019.
[28] S. W. C. Nikkelen, P. M. Valkenburg, M. Huizinga, and B. J. Bushman, "Media use and ADHD-related behaviors in children and adolescents: A meta-analysis." *Developmental Psychology*, vol. 50, no. 9, pp. 2228– 2241, sep 2014.
[29] M. D. Greicius, B. Krasnow, A. L. Reiss, and V. Menon, "Functional connectivity in the resting brain: A network analysis of the default mode hypothesis," *Proceedings of the National Academy of Sciences*, vol. 100, no. 1, pp. 253–258, dec 2002.
[30] B. Park, D. H. Han, and S. Roh, "Neurobiological findings related to internet use disorders," *Psychiatry and Clinical Neurosciences*, vol. 71, no. 7, pp. 467–478, sep 2016.
[31] J. A. Firth, J. Torous, and J. Firth, "Exploring the impact of internet use on memory and attention processes," *International Journal of Environmental Research and Public Health*, vol. 17, no. 24, p. 9481, dec 2020.
[32] S. Kühn and J. Gallinat, "Brains online: structural and functional correlates of habitual internet use," *Addiction Biology*, vol. 20, no. 2, pp. 415–422, feb 2014.
[33] K. K. Loh and R. Kanai, "Higher media multi-tasking activity is associated with smaller gray-matter density in the anterior cingulate cortex," *PLoS ONE*, vol. 9, no. 9, p. e106698, sep 2014.
[34] C. S. Green and D. Bavelier, "Action video game modifies visual selective attention," *Nature*, vol. 423, no. 6939, pp. 534–537, may 2003.
[35] J. C. Rosser, P. J. Lynch, L. Cuddihy, D. A. Gentile, J. Klonsky, and R. Merrell, "The impact of video games on training surgeons in the 21st century," *Archives of Surgery*, vol. 142, no. 2, pp. 181–186, feb 2007.
[36] A. F. Ward, "Supernormal: How the internet is changing our memories and our minds," *Psychological Inquiry*, vol. 24, no. 4, pp. 341–348, oct 2013.
[37] B. Sparrow, J. Liu, and D. M. Wegner, "Google effects on memory: Cognitive consequences of having information at our fingertips," *Science*, vol. 333, no. 6043, pp. 776–778, aug 2011.
[38] G. Dong and M. N. Potenza, "Behavioural and brain responses related to internet search and memory," *European Journal of Neuroscience*, vol. 42, no. 8, pp. 2546–2554, sep 2015.
[39] D. D. Giacomo, J. Ranieri, and P. Lacasa, "Digital learning as enhanced learning processing? cognitive evidence for new insight of smart learning," *Frontiers in Psychology*, vol. 8, aug 2017.
[40] A. Abbott, "Gaming improves multitasking skills," *Nature*, vol. 501, no. 7465, pp. 18–18, sep 2013.
[41] W. A. van der Schuur, S. E. Baumgartner, S. R. Sumter, and P. M. Valkenburg, "The consequences of media multitasking for youth: A review," *Computers in Human Behavior*, vol. 53, pp. 204–215, dec 2015.
[42] Y. T. Uhls, M. Michikyan, J. Morris, D. Garcia, G. W. Small, E. Zgourou, and P. M. Greenfield, "Five days at outdoor education camp without screens improves preteen skills with nonverbal emotion cues," *Computers in Human Behavior*, vol. 39, pp. 387–392, oct 2014.
[43] V. Bochicchio, N. M. Maldonato, P. Valerio, R. Vitelli, S. dell'Orco, and C. Scandurra, "A review on the effects of digital play on children's cognitive and socio-emotional development," in *2018 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, aug 2018.
[44] B.-C. Han, *The Burnout Society*, E. Butler, Ed. Stanford, California: Stanford Briefs, an imprint of Stanford University Press, 2015.
[45] A. A. Golroudbari and M. H. Sabour, "Recent advancements in deep learning applications and methods for autonomous navigation: A comprehensive review," 2023.
[46] D. C. Guastella and G. Muscato, "Learning-based methods of perception and navigation for ground vehicles in unstructured environments: A review," *Sensors*, vol. 21, no. 1, p. 73, dec 2020.
[47] E. Matheson, R. Minto, E. G. G. Zampieri, M. Faccio, and G. Rosati, "Human–robot collaboration in manufacturing applications: A review," *Robotics*, vol. 8, no. 4, p. 100, dec 2019.
[48] P. E. Dupont, B. J. Nelson, M. Goldfarb, B. Hannaford, A. Menciassi, M. K. O'Malley, N. Simaan, P. Valdastri, and G.-Z. Yang, "A decade retrospective of medical robotics research from 2010 to 2020," *Science Robotics*, vol. 6, no. 60, nov 2021.
[49] Y. Zhang and M. Lu, "A review of recent advancements in soft and flexible robots for medical applications," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 16, no. 3, mar 2020.
[50] M. K. Ho and T. L. Griffiths, "Cognitive science as a source of forward and inverse models of human decisions for robotics and control," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 5, no. 1, pp. 33–53, may 2022.
[51] X. Gao, Y. Wang, X. Chen, and S. Gao, "Interface, interaction, and intelligence in generalized brain–computer interfaces," *Trends in Cognitive Sciences*, vol. 25, no. 8, pp. 671–684, aug 2021.
[52] T. O. Zander and C. Kothe, "Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general," *Journal of neural engineering*, vol. 8, no. 2, p. 025 005, 2011.
[53] X.Tang,H.Shen,S.Zhao,N.Li,andJ.Liu,"Flexiblebrain–computer interfaces," *Nature Electronics*, vol. 6, no. 2, pp. 109–118, feb 2023.
[54] H. Kim, Y.-T. Kwon, H.-R. Lim, J.-H. Kim, Y.-S. Kim, and W.-H. Yeo, "Recent advances in wearable sensors and integrated functional devices for virtual and augmented reality applications," *Advanced Functional Materials*, vol. 31, no. 39, p. 2 005 692, oct 2020.
[55] T. Zhan, K. Yin, J. Xiong, Z. He, and S.-T. Wu, "Augmented reality and virtual reality displays: Perspectives and challenges," *iScience*, vol. 23, no. 8, p. 101 397, aug 2020.
[56] J. Carroll, L. Hopper, A. M. Farrelly, R. Lombard-Vance, P. D. Bamidis, and E. I. Konstantinidis, "A scoping review of augmented/ virtual reality health and wellbeing interventions for older adults: Redefining immersive virtual reality," *Frontiers in Virtual Reality*, vol. 2, jun 2021.
[57] P. M. G. Emmelkamp and K. Meyerbröker, "Virtual reality therapy in mental health," *Annual Review of Clinical Psychology*, vol. 17, no. 1, pp. 495–519, may 2021.
[58] I. Horváth, "Evolution of teaching roles and tasks in VR / AR-based education," in *2018 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, aug 2018, pp. 355–360.
[59] G. Makransky and G. B. Petersen, "Investigating the process of learning with desktop virtual reality: A structural equation modeling approach," *Computers & Education*, vol. 134, pp. 15–30, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360131519300259
[60] B. Schiavi, V. Havard, K. Beddiar, and D. Baudry, "BIM data flow architecture with AR/VR technologies: Use cases in architecture, engineering and construction," *Automation in Construction*, vol. 134, p. 104 054, feb 2022.
[61] S. Verma, L. Warrier, B. Bolia, and S. Mehta, "Past, present, and future of virtual tourism-a literature review," *International Journal of Information Management Data Insights*, vol. 2, no. 2, p. 100 085, nov 2022.
[62] J. Pirker, E. Loria, S. Safikhani, A. Kunz, and S. Rosmann, "Immersive virtual reality for virtual and digital twins: A literature review to identify state of the art and perspectives," in *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, mar 2022.
[63] T. Kaarlela, S. Pieska, and T. Pitkaaho, "Digital twin and virtual reality for safety training," in *2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, sep 2020.
[64] P. Galambos, C. Weidig, P. Baranyi, J. C. Aurich, B. Hamann, and O. Kreylos, "VirCA NET: A case study for collaboration in shared virtual space," in *2012 IEEE 3rd International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, dec 2012.
[65] I. L. Chamusca, C. V. Ferreira, T. B. Murari, A. L. Apolinario, and I. Winkler, "Towards sustainable virtual reality: Gathering design guidelines for intuitive authoring tools," *Sustainability*, vol. 15, no. 4, p. 2924, feb 2023.
[66] T. Setti and A. B. Csapo, "A canonical set of operations for editing dashboard layouts in virtual reality," *Frontiers in Computer Science*, vol. 3, jul 2021.
[67] I. Horváth and Á. B. Csapó, "Motivations and tools relevant to personalized workspaces in VR environments," *Electronics*, vol. 12, no. 9, apr 2023. [Online]. Available: https://www.mdpi.com/2079-9292/12/9/2059

[68] K. Shah, A. Salunke, S. Dongare, and K. Antala, "Recommender systems: An overview of different approaches to recommendations," in *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*. IEEE, mar 2017.

[69] J. M. Valdez Mendia and J. J. A. Flores-Cuautle, "Toward customer hyper-personalization experience — a data-driven approach," *Cogent Business & Management*, vol. 9, no. 1, Feb. 2022.

[70] J. Hanson, M. Wei, S. Veys, M. Kugler, L. Strahilevitz, and B. Ur, "Taking data out of context to hyper-personalize ads: Crowdworkers' privacy perceptions and decisions to disclose private information," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI '20. ACM, Apr. 2020.

[71] M. Jesse and D. Jannach, "Digital nudging with recommender systems: Survey and future directions," *Computers in Human Behavior Reports*, vol. 3, p. 100052, jan 2021.

[72] S. Russell, *Human Compatible: Artificial Intelligence and the Problem of Control*. Penguin, 2019.

[73] V. Dörfler, *What Every CEO Should Know About AI*. University Press, Mar. 2022. Cambridge

[74] E.Adamopoulou and L. Moussiades, "A noverview of chatbot technology," in *IFIP Advances in Information and Communication Technology*. Springer International Publishing, 2020, pp. 373–383.

[75] A. Rapp, L. Curti, and A. Boldi, "The human side of human-chatbot interaction: A systematic literature review of ten years of research on text-based chatbots," *International Journal of Human-Computer Studies*, vol. 151, p. 102630, jul 2021.

[76] A. Sheth, H. Y. Yip, A. Iyengar, and P. Tepper, "Cognitive services and intelligent chatbots: Current perspectives and special issue introduction," *IEEE Internet Computing*, vol. 23, no. 2, pp. 6–12, mar 2019.

[77] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[78] S. Bubeck, V. Chandrasekaran, R. Eldan, J. Gehrke, E. Horvitz, E. Kamar, P. Lee, Y. T. Lee, Y. Li, S. Lundberg, H. Nori, H. Palangi, M. T. Ribeiro, and Y. Zhang, "Sparks of artificial general intelligence: Early experiments with GPT-4," 2023.

[79] T. Ullman, "Large language models fail on trivial alterations to theory-of-mind tasks," 2023.

[80] N. Shapira, M. Levy, S. H. Alavi, X. Zhou, Y. Choi, Y. Goldberg, M. Sap, and V. Shwartz, "Clever hans or neural theory of mind? stress testing social reasoning in large language models," 2023.

[81] M. Kosinski, "Theory of mind may have spontaneously emerged in large language models," 2023.

[82] G. Stojnić, K. Gandhi, S. Yasuda, B. M. Lake, and M. R. Dillon, "Commonsense psychology in human infants and machines," *Cognition*, vol. 235, p. 105406, jun 2023.

[83] P. Baranyi, Á. Csapó, T. Budai, and G. Wersényi, "Introducing the concept of internet of digital reality – part i," *Acta Polytechnica Hungarica*, vol. 18, no. 7, pp. 225–240, 2021.

[84] G. Wersényi, Á. Csapó, T. Budai, and P. Baranyi, "Internet of digital reality: Infrastructural background – part ii," *Acta Polytechnica Hungarica*, vol. 18, no. 8, pp. 91–104, 2021.

[85] A. S.-K. Pang, *The distraction addiction: Getting the information you need and the communication you want, without enraging your family, annoying your colleagues, and destroying your soul*. Hachette UK, 2013.

[86] P. Baranyi and A. B. Csapo, "Revisiting the concept of generation CE - generation of cognitive entities," in *2015 6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, oct 2015, pp. 583–586.

[87] J. Searle, "I married a computer," *The New York Review of Books*, vol. 8, 1999.

[88] N. A. of Sciences, In the Light of Evolution: Volume VII: The Human Mental Machinery, C. J. Cela-Conde, R. G. Lombardo, J. C. Avise, and F. J. Ayala, Eds. Washington, DC: *The National Academies Press*, 2014. [Online]. Available: https://nap.nationalacademies.org/catalog/18573/in-the-light-of-evolution-volume-vii-the-human-mental

[89] B. Berki, "Overview of the relationship between humanspatialabilities and GPS usage," in *2022 IEEE 1st International Conference on Cognitive Mobility (CogMob)*. IEEE, Oct. 2022, pp. 21–24.

[90] Á. B. Csapó and P. Z. Baranyi, "A unified terminology for the structure and semantics of CogInfoCom channels," *Acta Polytechnica Hungarica*, vol. 9, no. 1, pp. 85–105, 2012.

[91] P. Galambos, "Vibrotactile feedback for haptics and telemanipulation: Survey, concept and experiment," *Acta Polytechnica Hungarica*, vol. 9, no. 1, pp. 41–65, 2012.

[92] T. Hildebrandt, T. Hermann, and S. Rinderle-Ma, "A sonification system for process monitoring as secondary task," in *2014 5th IEEE Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, nov 2014.

[93] J. Yang and T. Hermann, "A zen garden interface for the interactive control of sonic ambiences in smart environment," in *2015 6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, oct 2015.

[94] J. Hammerschmidt, T. Hermann, A. Walender, and N. Kromker, "InfoPlant: Multimodal augmentation of plants for enhanced human-computer interaction," in *2015 6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, oct 2015.

[95] B. Endrstova, M. Macik, and L. Treml, "Reprobooktor: A concept of audiobook player for visually impaired older adults," in *2018 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, aug 2018.

[96] M. Macik, V. Gintner, D. Palivcova, and I. Maly, "Tactile symbols for visually impaired older adults," in *2018 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, aug 2018.

[97] A. Csapo and P. Baranyi, "The spiral discovery method: An interpretable tuning model for CogInfoCom channels," *Journal of Advanced Computational Intelligence*, vol. 16, no. 2, pp. 358–367, mar 2012.

[98] P. Baranyi, A. Csapo, and P. Varlaki, "An overview of research trends in CogInfoCom," in *IEEE 18th International Conference on Intelligent Engineering Systems INES 2014*. IEEE, jul 2014.

[99] W.-K. Leung, K.-W. Yuen, K.-H. Wong, and H. Meng, "Development of text-to-audiovisual speech synthesis to support interactive language learning on a mobile device," in *2013 IEEE 4th International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, dec 2013.

[100] G. Wirsching and R. Lorenz, "Towards meaning-oriented language modeling," in *2013 IEEE 4th International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, dec 2013.

[101] B. J. Kroger, T. Bekolay, and C. Eliasmith, "Modeling speech production using the neural engineering framework," in *2014 5th IEEE Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, nov 2014.

[102] D. Sztahó, M. G. Tulics, K. Vicsi, and I. Valálik, "Automatic estimation of severity of parkinson's disease based on speech rhythm related features," in *2017 8th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, 2017, pp. 000 011–000 016.

[103] S. Ondas, J. Juhar, E. Kiktova, and J. Zimmermann, "Anticipation in speech-based human-machine interfaces," in *2018 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, aug 2018.

[104] A. Esposito, G. Raimo, M. Maldonato, C. Vogel, M. Conson, and G. Cordasco, "Behavioral sentiment analysis of depressive states," in *2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, sep 2020.

[105] L. McLaren, M. Koutsombogera, and C. Vogel, "Gaze, dominance and dialogue role in the MULTISIMO corpus," in *2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, sep 2020.

[106] D. Sztaho, A. Orosz, and I. Valalik, "Articulation correctness measurement of parkinson's disease using low resource-intensitve segmentation methods," in *2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, sep 2020.

[107] M. Gnjatovic and N. Macek, "An entropy minimization approach to dialogue segmentation," in *2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, sep 2020.

[108] C. Navarretta, "Speech pauses and pronominal anaphors," *Frontiers in Computer Science*, vol. 3, may 2021.

[109] R. Römer, P. beim Graben, M. Huber-Liebl, and M. Wolff, "Unifying physical interaction, linguistic communication, and language acquisition of cognitive agents by minimalist grammars," *Frontiers in Computer Science*, vol. 4, jan 2022.

[110] C. Vogel, M. Koutsombogera, and J. Reverdy, "Aspects of dynamics in dialogue collaboration," *Electronics*, vol. 12, no. 10, p. 2210, may 2023.

[111] L. J. Laki and Z. G. Yang, "Sentiment analysis with neural models for hungarian,"*Acta Polytechnica Hungarica*, vol. 20, no.5, pp. 109–128, 2023.

[112] D. Geszten, A. Komlódi, K. Hercegfi, B. Hámornik, A. Young, M. Köles, and W. G. Lutters, "A content-analysis approach for exploring usability problems in a collaborative virtual environment," *Acta Polytechnica Hungarica*, vol. 15, no. 5, pp. 67–88, 2018.

[113] D. Geszten, B. P. Hámornik, and K. Hercegfi, "Finding treasure in the wild: A field study of a collaborative design software," in *Accentuated Innovations in Cognitive Info-Communication*. Springer International Publishing, sep 2022, pp. 185–202.

[114] B. Wortelen, S. Feuerstack, M.-C. Harre, S. Senturk, F. You, and J. Wang, "Engineering optimized control panel designs for electric driving," in *2019 10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, oct 2019.

[115] V. Szucs, C. Sik-Lanyi, and T. Guzsvinecz, "Presenting the user's focus in needs & development (UFIND) method and its comparison to other design methods," in *2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, sep 2020.

[116] N. Vittayaphorn, G. Lohaburananont, J. Bhumtakwong, K. Udompongsanont, M. Uchida, N. Asavarojkul, P. Rodphol, P. Sajjapong, R. Boonribsong, S. Chanthateyanonth, T. Julerttrakul, and A. Pyae, "Design and development of a user-centered mobile application for intermodal public transit in Bangkok: A design thinking approach," *Infocommunications Journal*, vol. 15, no. Special Issue, pp. 41–52, 2023.

[117] A. Torok, "From human-computer interaction to cognitive infocommunications: A cognitive science perspective," in *2016 7th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, oct 2016, pp. 433–438.

[118] A. M. Putnoki, D. Mattyasovszky-Philipp, and B. Molnár, "Cognitive infocommunication and carbon-carbon interaction," in *2022 13th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, sep 2022.

[119] B. Molnár and D. Mattyasovszky-Philipp, "Cognitive resonance and the architecture issues of cognitive information systems," in *Accentuated Innovations in Cognitive Info-Communication*. Springer International Publishing, sep 2022, pp. 29–56.

[120] A. Ito and K. Nakada, "UI design based on traditional japanese gesture," in *2019 10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, oct 2019.

[121] Y. Muramatsu, M. Niitsuma, and T. Thomessen, "Perception of tactile sensation using vibrotactile glove interface," in *2012 IEEE 3rd International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, dec 2012.

[122] P. Zikmund, M. Macik, L. Dubnicky, and M. Horpatzska, "Comparison of joystick guidance methods," in *2019 10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, oct 2019.

[123] Z. Mikovec, P. Slavik, and J. Zara, "Cultural heritage, user interfaces and serious games at CTU prague," in *2009 15th International Conference on Virtual Systems and Multimedia*. IEEE, sep 2009.

[124] J. Jamsa and H. Kaartinen, "Adaptive user interface for assisting the drivers' decision making," in *2015 6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, oct 2015.

[125] I. Péntek and A. Adamkó, "Monitoring pandemics, using sensor data from smart ecosystems," *Acta Polytechnica Hungarica*, vol. 20, no. 5, pp. 75–90, 2023.

[126] G. Sallai, "Evolution of digitization toward the internet of digital & cognitive realities and smart ecosystems," *Infocommunications Journal*, vol. 15, no. Special Issue, pp. 2–8, 2023.

[127] L. Blázovics, B. Forstner, H. Charaf, and F. H. P. Fitzek, "On the benefits of cognitive infocommunication for mobile communication nodes using cooperative concepts," in *2013 IEEE 4th International Conference on Cognitive Infocommunications (CogInfoCom)*, Dec 2013, pp. 693–698.

[128] K. Csorba, P. Ekler, A. Bori, and H. Charaf, "Analysis of mobile bittorrent client behavior," in *2013 IEEE 4th International Conference on Cognitive Infocommunications (CogInfoCom)*, Dec 2013, pp. 613– 618.

[129] P. Baranyi and A. Gilányi, "Mathability: Emulating and enhancing human mathematical capabilities," in *2013 IEEE 4th international conference on cognitive Infocommunications (CogInfoCom)*. IEEE, 2013, pp. 555–558.

[130] K. Chmielewska, "From mathability to learnability," in *2019 10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, oct 2019.

[131] M. Török, M. J. Tóth, and A. Szöllősi, "Foundations and perspectives of mathability in relation to the coginfocom domain," in *2013 IEEE 4th International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, 2013, pp. 869–872.

[132] K. Chmielewska and A. Gilnyi, "Educational context of mathability," *Acta Polytechnica Hungarica*, vol. 15, no. 5, 2018.

[133] K. Bubnó and V. L. Takács, "Cognitive aspects of mathematics-aided computer science teaching," *Acta Polytechnica Hungarica*, vol. 16, no. 5, aug 2019.

[134] Z. Kvasznicza, I. Gyurcsek, G. Elmer, V. Bagdán, and I. Horváth, "Mathability of EMC emission testing for mission crucial devices in GTEM waveguide," *Acta Polytechnica Hungarica*, vol. 18, no. 1, pp. 159–173, 2021.

[135] K. Chmielewska, A. Lukasiewicz, K. Mroczynska, and A. Gilányi, "Mathematical and technical education – experiences gained during online teaching," in *2022 13th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, sep 2022.

[136] K. Bubnó and V. L. Takács, "Mathematical and computational awareness before and after the pandemic," *Frontiers in Education*, vol. 7, sep 2022.

[137] I. Horváth, Á. B. Csapó, B. Berki, A. Sudár, and P. Baranyi, "Definition, background and research perspectives behind 'Cognitive Aspects of Virtual Reality' (cVR)," *Infocommunications Journal*, vol. 15, no. Special Issue, pp. 9–14, 2023.

[138] S. Korecko, M. Hudak, B. Sobota, M. Marko, B. Cimrova, I. Farkas, and R. Rosipal, "Assessment and training of visuospatial cognitive functions in virtual reality: proposal and perspective," in *2018 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, aug 2018.

[139] Z. G. Török and Á. Török, "Looking at the map – or navigating in a virtual city: interaction of visuospatial display and spatial strategies in VR," in *2018 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, aug 2018, pp. 327–332.

[140] A. Sudár and Á. Csapó, "Interaction patterns of spatial navigation and smartboard use in VR workspaces," in *Accentuated Innovations in Cognitive Info-Communication*. Springer International Publishing, sep 2022, pp. 149–166.

[141] B. Berki, "Role of presence, memory and spatial ability in a desktop virtual reality," in *Accentuated Innovations in Cognitive Info-Communication*. Springer International Publishing, sep 2022, pp. 79– 97.

[142] A. Sudár and Á. B. Csapó, "Descriptive markers for the cognitive profiling of desktop 3D spaces," *Electronics*, vol. 12, no. 2, p. 448, jan 2023.

[143] ——, "Elicitation of content layout preferences in virtual 3D spaces based on a free layout creation task," *Electronics*, vol. 12, no. 9, apr 2023. [Online]. Available: https://www.mdpi.com/2079-9292/12/9/2078

[144] F. Erdős, R. Németh, and B. Firuza, "Virtual teamwork in gamified 3D environment," *Infocommunications Journal*, vol. 15, no. Special Issue, pp. 15–20, 2023.

[145] A. Kovari, "CogInfoCom supported education: A review of CogInfoCom based conference papers," in *2018 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, IEEE. IEEE, aug 2018, pp. 000 233–000 236.

[146] M. Dergham and A. Gilányi, "Application of virtual reality in kinematics education," in *Proceedings of the 10th IEEE International Conference on Cognitive Infocommunications: CogInfoCom* 2019, P. Baranyi, Ed. IEEE, 2019, pp. 107–112.

[147] G. Molnár and D. Sík, "Smart devices, smart environments, smart students – a review on educational opportunities in virtual and aug- mented reality learning environments," in *2019 10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, oct 2019, pp. 495–498.

[148] T. A. Mikropoulos and J. Bellou, "The unique features of educational virtual environments," in *Proceedings of the IADIS International Conference on e-Society*, P. Isáıas, M. McPherson, and F. Bannister, Eds. Dublin, Ireland: IADIS Press, Jul. 2006, pp. 122–128.

[149] T. Budai and M. Kuczmann, "A multi-purpose virtual laboratory with interactive knowledge integration," in *Proceedings of the 10th IEEE International Conference on Cognitive Infocommunications: CogInfoCom* 2019, P. Baranyi, Ed. IEEE, 2019, pp. 529–532.

[150] C. H. Wijkmark, I. Heldal, S. Fankvist, and M.-M. Metallinou, “Remote virtual simulation for incident commanders: Opportunities and possibilities,” in *2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, sep 2020.
[151] I. Horváth,“Experiences and guidelines for developing 3D VR curricula,” in *Accentuated Innovations in Cognitive Info-Communication*. Springer International Publishing, sep 2022, pp. 99–123.
[152] I. K. Boda and E. Tóth, “Text-based second language learning in the three-dimensional space,” in *Accentuated Innovations in Cognitive Info-Communication*. Springer International Publishing, sep 2022, pp. 125–148.
[153] T. Guzsvinecz, É. Orbán-Mihálykó, E. Perge, and C. Sik-Lányi, “The effect of engineering education, on spatial ability, in virtual environments,” *Acta Polytechnica Hungarica*, vol. 20, no. 5, pp. 11–30, 2023.
[154] B. Sobota, Š. Korečko, and M. Mattová, “Web-based 3D virtual environments utilization in primary and secondary education of children with multiple impairments,” *Electronics*, vol. 12, no. 13, p. 2792, jun 2023.
[155] H. V. Kerrebroeck, M. Brengman, and K. Willems, “When brands come to life: experimental research on the vividness effect of virtual reality in transformational marketing communications,” *Virtual Reality*, vol. 21, no. 4, pp. 177–191, jan 2017.
[156] B. Berki, “2D advertising in 3D virtual spaces,” *Acta Polytechnica Hungarica*, vol. 15, no. 3, pp. 175–190, 2018.
[157] K. Boros, G. Bujdoso, C. M. Novac, and O. C. Novac, “E-health promotion virtual reality services in MaxWhere VR spaces – design and development,” in *2019 10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, oct 2019.
[158] S. R. van der Kruk, R. Zielinski, H. MacDougall, D. Hughes-Barton, and K. M. Gunn, “Virtual reality as a patient education tool in healthcare: A scoping review,” *Patient Education and Counseling*, feb 2022.
[159] F. Bellalouna, “Industrial use cases for augmented reality application,” in *2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, sep 2020.
[160] M. Luimula, M. Linder, S. Pieska, E. Laimio, T. Lahde, and P. Porramo, “Unlimited safety productivity – a finnish perspective using virtual learning methods to improve quality and productivity in the construction industry,” in *2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, sep 2020.
[161] M. Luimula, E. Markopoulos, J. K. Kaakinen, P. Markopoulos, N. Laivuori, and W. Ravyse, “Eye tracking in maritime immersive safe oceans technology,” in *2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, sep 2020.
[162] N. Nascivera, Y. M. Alfano, T. Annunziato, M. Messina, V. S. Iorio, V. Cioffi, R. Sperandeo, M. Rosato, T. Longobardi, and N. M. Maldonato, “Virtual empathy : The added value of virtual reality in psychotherapy,” in *2018 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, aug 2018.
[163] A. Tafuro, G. Masi, M. Terracciano, N. M. Maldonato, M. Bottone, R. Sperandeo, C. Scandurra, and B. Muzii, “Virtual reality in the treatment of social anxiety disorder: A systematic review,” in *2022 13th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, sep 2022.
[164] M. Zöldy and P. Baranyi, “Cognitive mobility – CogMob,”in *12 th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, 2021, pp. 921–925.
[165] M. Zöldy and P. Baranyi, “The cognitive mobility concept,” *Infocommunications Journal*, vol. 15, no. Special Issue, pp. 35–40, 2023. 12
[166] A. Kő, I. Szabó, Á. B. Csapó, T. Kovács, L. Lőrincz, and P. Baranyi, “Digital & Cognitive Corporate Reality,” *Infocommunications Journal*, vol. 15, no. Special Issue, pp. 2–10, 2023.
[167] G. Kampis, L. Gulyas, and S. Soos, “Do institutions think? comparing past success and future investment,” in *2013 IEEE 4th International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, dec 2013.
[168] D. Susniene, A. Torma, A. Buruzs, M. F. Hatwagner, and L. T. Koczy, “Using fuzzy cognitive map approach to model the casual relationships in stakeholder management at companies,” in *2014 5th IEEE Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, nov 2014.
[169] P. Waldbuesser and L. I. Komlosi, “Empirical findings in cognitive entity management: A challenge in the digital era,” in *2015 6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, oct 2015.
[170] P. Szmodics, “Knowledge-based process management,” in *2015 6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, oct 2015.
[171] D. Rumiński, M. Maik, and K. Walczak, “Visualizing financial stock data within an augmented reality trading environment,” *Acta Polytechnica Hungarica*, vol. 16, no. 6, pp. 223–239, 2019.

Péter Baranyi established the Cognitive Infocommunications concept around 2010. It is a scientific discipline today focusing on the new cognitive capabilities of the blended combination of human and informatics. It has an annual IEEE International Conference and a number of scientific journal special issues. He invented the TP model transformation which is a higher-order singular value decomposition of continuous functions. It has a crucial role in non-linear control design theories and opens new ways for optimization. He is the inventor of MaxWhere which is the first 3D platform including 3D web, 3D browser, 3D store, and 3D Cloud. His research group published a number of journal papers firstly reporting that users get 40-50% better effectiveness in 3D digital environments. These results got a very high international impact.

Borbála Berki obtained her PhD degree at Széchenyi István University in 2023. From October 2022 to October 2023, she served as an assistant lecturer at Óbuda University. Currently, she holds the positions of junior research fellow and assistant professor at Corvinus University of Budapest. Dr. Berki’s main research interests lie in developing experimental methods for studying users’ cognitive capabilities towards e.g. comprehension and memory in 2D versus 3D digital environments. Even as a young researcher having recently obtained her PhD degree, Dr. Berki’s work has garnered over 290 independent citations in the scientific literature.

Ádám B. Csapó obtained his PhD degree at the Budapest University of Technology and Economics in 2014. From 2016 to 2022, he has worked as an associate professor at the Széchenyi István University in Győr; and, since 2022, he continued his academic journey as as an associate professor at Óbuda University, Budapest, Hungary. Since September 2023, he has held the position of associate professor at Corvinus University of Budapest. Dr. Csapó’s research focuses on cognitive infocommunication channels in virtual collaboration environments, i.e. enabling users to communicate with each other and with their spatial surroundings in novel and effective ways. At the same time, he has been involved in the development of assistive technology for the visually impaired, as well as in the development of a commercial VR platform. Dr. Csapó has over 50 publications, including 1 co-authored book and 20 journal papers, and has actively participated in the organization of numerous international conferences and special issues.

Computation of Accessibility Score of Educational Institute Webpages using Machine Learning Approaches

Jinat Ara, and Cecilia Sik-Lanyi

***Abstract*—The availability of digital platforms by ensuring accessibility and usability is considered a virtual gateway to provide a wide array of information to its stakeholders. An accessible web platform can disseminate information among a variety of target audiences. Thereby accessibility of academic web pages requires special attention. Herein we proposed an accessibility computation approach for higher education institute webpage (Homepage) in the context of universities in Hungary. The proposed approach incorporated two machine learning (ML) classifiers: Random Forest (RF), and Decision Tree (DT) to experiment on our custom dataset to compute the overall accessibility score. Performance of ML methods validated through confusion matrix and classification report result. The empirical results of ML methods and statistical evaluation showed poor accessibility scores which depicts that none of the selected web pages are free from accessibility issues associated with disabilities. As such, accessibility is a crucial aspect that needs further concern as most of the considered academic webpages have experienced accessibility issues and showed improvement demands.**

***Index Terms*—Accessibility validation; machine learning methods; questionnaire analysis; academic institute.**

I. Introduction

The growth of web technologies brought immense progress in different spheres of life, especially in accessing digital information [1][2] from various platforms, such as the healthcare sector, banking sector, education sector, and others [3]. In this modern era, focusing on any sector, having a digital platform (e.g., webpages) for providing information is inevitable. In many cases, web pages act as the primary resources, especially for communicating with various stakeholders. Therefore, different stakeholders access web pages frequently for their required information. For example, from the perspective of the university, academic and prospective students might need to download their course curriculum, class schedule, campus news, admission requirements, and other information. Thus, digital platforms or webpages should be accessible to serve their stakeholders to compete globally. However, the emerging concern is that most web pages are not developed concerning the accessibility perspective [4].

The authors are with the Department of Electrical Engineering and Information System, University of Pannonia, Veszprem, Hungary.
E-mail: {jinat.ara@mik.uni-pannon.hu, lanyi.cecilia@mik.uni-pannon.hu).

These developments limit the universal inclusion of potential users. Accessible web generally refers to the design and development of websites (multiple web pages) in a manner that is effective for people with disabilities and without disabilities [5]. Nowadays, web designers and developers are trying to incorporate several complex functionalities (e.g., dynamic, drop-down menu) and components (e.g., images, videos) into their web pages to make them more interactive. Though these interactive functionalities are prominent to attract more people, they limit the accessibility concept for users with disabilities [6] [7]. Therefore, it is increasingly important to design and develop web pages considering accessibility manners and follow accessibility practices to serve equal access to resources. Concerning educational institute webpages, the introduction of accessible university webpages is not only beneficial for students with impairments but also for the university authorities and other associate practitioners for their academic progress. From the experience of the COVID-19 pandemic, unlike in other countries, the importance of improving the quality of life for students with disabilities at higher education levels has increased in Hungary. Despite addressing numerous challenges faced by students with disabilities in past studies, the accessibility aspects for higher education students have been overlooked from the beginning. Thus, the consequence of the current accessibility limitation of higher education webpages has grown dramatically, particularly university webpages.

Over the years, researchers from many countries have conducted accessibility evaluations of web pages concerning issues and benefits of people with disabilities at different levels. The record of past literature depicts that the majority of the proposed approach evaluated the effectiveness of the webpages concerning their quality (e.g., broken link, interactivity) and usability (e.g., HTML page, aesthetic, design, page size) [8-10]. Concerning the increasing number of people with disabilities, researchers from different backgrounds (e.g., technology, education, neurodevelopment) have sought to evaluate the accessibility of online platforms [11]. Some researcher contributed their effort to develop an effective approach to validate the web, identify accessibility issues, and compute accessibility barrier scores. Among several approaches, fuzzy inference-based evaluation, regression model development, and variable magnitude approaches are prominent. Besides, nowadays several automated web accessibility testing tools have been developed that provide interactive accessibility

DOI: 10.36244/ICJ.2024.5.6

reports about the tested website. However, these approaches and existing automated tools incorporate some specific attributes of websites according to the Web Content Accessibility Guidelines (WCAG) to evaluate their accessibility status. Though the latest version of WCAG 2.2 [12] is a complete guideline for accessibility features, it has limited consideration about some issues with people with disabilities such as whether the website is active or deactivated, website has a manual text size adjustment option, manual font family adjustment option, manual color adjustment option, user information requirement, CAPTCHA issues, usefulness of internal/external links, used images, inserted video and audio content. As these features are not directly possible to evaluate in an automatic manner, WCAG does not provide a clear indication about these aspects though without considering these aspects, it is not possible to ensure complete accessibility of the developed websites.

In that manner, our prime focus is to evaluate websites considering the above-mentioned aspects and compute the overall accessibility score based on these aspects. The prime challenges of this work related to the prepared dataset as there is no dataset has been found that evaluated websites according to these ten aspects. Thus, to conduct this work, we used our custom dataset that we prepared according to the ten aspects mentioned previously and incorporated two machine learning approaches (Random Forest (RF), and Decision tree (DT)) that might be effective to validate the accessibility of the web platform. Besides, none of the work showed their contribution incorporating machine learning (ML) classifiers to add the benefits in this particular domain, such as accessibility of the digital platform. The classification result of ML methods has been evaluated through classification accuracy, precision, sensitivity, specificity, and F1 score. Overall, this paper aims to contribute to web accessibility research by proposing an accessibility score computation system to validate the accessibility of university web pages. The proposed system is platform-independent and dynamic, thus can evaluate and perform a comparative analysis of any academic webpages of a different country.

The remainder of this paper is structured as follows. Section 2 presents a brief discussion of the background and related work about web accessibility. Section 3 presents the methods and materials of the proposed framework, including dataset preparation and description, and system architecture and design. Section 4 presents the experimental analysis including classifiers and website performance analysis. Section 5 presents a detailed discussion. Finally, the paper is concluded through conclusion and future work.

II. RELATED STUDIES

Higher education institutions (e.g., universities) play a vital role in developing society and in the scientific community by educating the young mind. For example, university websites (multiple web pages) are responsible for providing information to the community. As such evaluating webpages, researchers considered several methods to identify the quality of the webpage to ensure the inclusion of stakeholders with disabilities. Numerous past studies focused on the education domain to evaluate their accessibility. For example, Chopra et al. [13] enhanced the importance of e-learning at the higher education level. They conducted a questionnaire-based statistical evaluation of university websites. They added that website quality, service quality, and information quality are leading factors to influence user satisfaction and net benefits, which is crucial to consider during web development. Mittal et al. [14] proposed a website quality evaluation process using fuzzy logic/technique. They incorporated fuzzy logic to assess websites in terms of several metrics such as loading time, response time, mark-up validation, broken link, accessibility error, size, page rank, frequency of update, traffic, and design optimization. In another work, Malhotra et al. [15] focused on website quality prediction through an automated Web Metrics Analyzer tool called Neuro-fuzzy inference models. This study confirms that a fuzzy logic-based website analyzer is feasible for predicting the quality of the website.

Further, few studies proposed several frameworks to evaluate academic or higher institute websites. As such, Rashida et al. [16] developed an automated web-based tool to investigate university websites following the content of information, loading time, and overall performance. Their result shows that most university websites did not meet users' satisfaction. Olaleye et al. [17] proposed a framework called WebFUQII based on the web analytical tools WebQual and SITEQUAL, considering ease of use, processing speed, aesthetic design, interactive responsiveness, entertainment, and trust and usefulness.

Concerning accessibility, Alahmadi [18] proposed a multi-model accessibility evaluation framework for university websites for deaf, visually impaired, and Deaf-blindness students. They incorporated automated tools to generate accessibility reports, source code mining to evaluate media content accessibility errors, and human evaluation to validate the assessment result. Focusing on particular disability types, such as people with vision impairment, Hassouna et al. [19] proposed a framework incorporating manual assessment (user and expert testing) and automatic assessment (Cynthiasays) to evaluate the accessibility of university websites.

Following this, few studies considered only automatic accessibility testing tools to investigate the accessibility of higher institute websites. For example, Verkijika et al. [20] evaluated 26 South African university websites through two automated validators (e.g., AChecker and TAW). This investigation reveals that the appearance of broken links and failing Google Mobile-friendly Test is the frequent issue that leads to accessibility issues. AlMeraj et al. [21] evaluated the accessibility of 41 higher education institution websites in Kuwait. They validated considering several automatic accessibility testing tools, specifically AChecker, Total Validator, MAUVE++, WAVE, and HTML/CSS/ARIA. This evaluation concludes that during the website design and development, accessibility is not accounted which leads to an increased number of inaccessible websites.

Nowadays, web researchers have shown their active participation in considering machine learning methods to evaluate the quality and usability of websites. For example, Dhiman et al. [22] proposed the most recent work that focused on machine learning methods to evaluate the performance of tools to identify website quality. They have implemented logistic regression and six machine learning methods, such as Random Forest, Adaboost, Bagging, Multilayer Perceptron, and Bayes Net. They depict that the machine learning model is more effective in evaluating website quality than other approaches. Addressing this research work, we aim to extend our evaluation by focusing on accessibility validation of the higher institute websites using ML techniques focusing on the eight aspects related to the website features. To evaluate the accessibility, we considered top university web pages in Hungary.

III. Material and Methods

This work has focused on webpage accessibility evaluation incorporating two machine learning classifiers. The prime objective is to observe the performance of the selected ML classifier to identify its effectiveness and then evaluate website accessibility according to the classification result. Figure 1 shows the working diagram that illustrates the materials and methods used in this study.

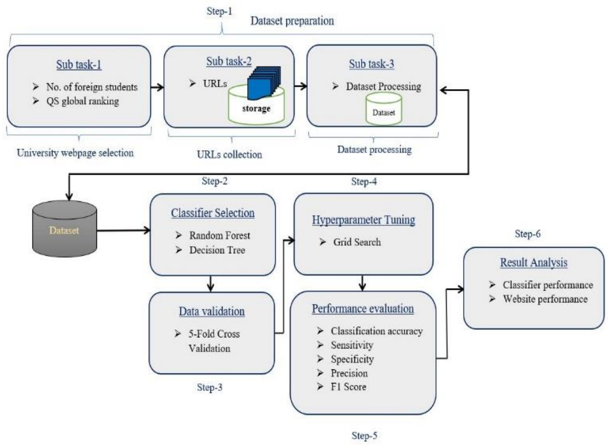


Fig.1. The System Architecture of the proposed model

A. Dataset Preparation

Dataset preparation is represented through multiple sub-tasks to get university webpage information to evaluate their quality and accessibility. University webpage selection, URL collection, and Dataset processing are the three sub-tasks of dataset preparation, as shown in Figure 1.

a) University Webpage Selection: The University webpage is the primary resource for a wide array of information such as departments, subjects, tuition fees, faculty, research groups, scholarships, etc., that help prospective students in their university selection. To investigate the quality and accessibility, initially, university selection is an important and complex task. To conduct this task, we considered Webometrics university ranking [23] for selecting the top five universities in Hungary. Then, we ranked the top five universities according to their number of international students and QS world ranking, as shown in Table 1.

TABLE I
DEMOGRAPHIC INFORMATION OF THE SELECTED WEBSITES

Websites	URLs	No. of foreign students	QS ranking
Web-1	https://u-szeged.hu/english	5000	551
Web-2	https://unideb.hu/en	4000	600
Web-3	https://www.elte.hu/en/	3000	700
Web-4	http://www.bme.hu/?language=en	1900	801
Web-5	https://www.ceu.edu/	962	124 (by subject)

b) Collection of URLs: The Uniform Resource Locator (URL) is a valuable resource for information extraction from the university website. For the selected top five universities, we stored their homepage URLs. Table 1 shows the university list with its URLs.

c) Dataset Processing: For dataset processing, we conducted a preliminary survey to understand the importance of our considered ten (10) aspects (availability, manual text size adjustment option, manual font family adjustment option, manual color adjustment option, user information requirement, CAPTCHA, usefulness of internal/external links, images, inserted video, and audio content) in terms of their effectiveness to the people with disability to represent the accessibility of the websites. Six users participated in this survey including vision problems (4), and cognitive problems (2) as the selected eight features are more likely to cause difficulty to these groups of people with disabilities. All the participants are active on the Internet platform for their professional work and daily activities and are between 25-50 years old. In the online survey, we asked participants to provide their opinions about whether these ten aspects are useful for them to understand the website content effectively. All the users expressed their positive opinion that these aspects are useful for understanding the web content and can improve their experience in internet platform browsing.

According to the positive feedback from the preliminary survey, we prepared our dataset incorporating human observation where 23 participants observed the selected five websites according to the selected ten aspects. All participants were university students from the Department of Electrical Engineering and Information Systems, at the University of Pannonia, Hungary. All the participants had adequate knowledge of the field of interactive design and web development. To evaluate the selected websites, we arranged online participation where we shared our prepared 10 questions related to the considered ten aspects (shown in Table 2) and website resources that need to be observed. All the participants observed the websites and answered each of the questions according to their understanding. After obtaining the responses from users for five selected websites, we labeled the responses in terms of ‘Accessible’, ‘Partially Accessible’, and ‘Not Accessible’ metrics where for all positive responses or ten positive responses, we labeled them as "Accessible"; for having >6 negative response (out of 10), we labeled it as "Not accessible" and rest of the responses having =<6 negative response (out of 10) were labeled as "Partially Accessible". Figure 2 shows the dataset preparation flowchart to represent the entire process in detail. In total 23 responses were recorded to each dataset related to eight features of the website and classified into three levels of accessibility status. However, the five tested websites' observation results have been incorporated as labeled datasets in our system to conduct the experimental analysis. As our data are categorical, we followed LabelEncoding to encode the data (label/categories). LabelEncoding is a popular categorical data encoding process. We used LabelEncoding through the sklearn LabelEncoder () function.

TABLE II
SELECTED QUESTIONS FOR THE PRELIMINARY SURVEY

Questions	**Response**
Question-1: Is the webpage available?	Yes/No
Question 2: Does the webpage have a manual text size adjustment option?	Yes/No
Question 3: Does the webpage have a manual font family adjustment option?	Yes/No
Question 4: Does the webpage have a manual color adjustment option?	Yes/No
Question 5: Does the webpage require user information?	Yes/No
Question 6: Does the webpage require CAPTCHA?	Yes/No
Question 7: Do the internal/external links are useful?	Yes/No
Question 8: Does the webpage images are useful?	Yes/No
Question 9: Is the webpage video content useful?	Yes/No
Question 10: Is the webpage audio content useful?	Yes/No

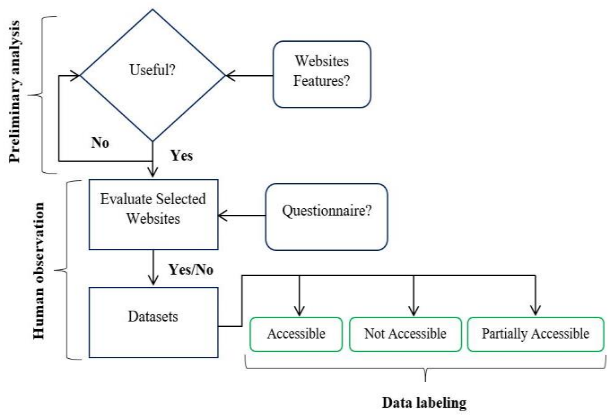


Fig. 2. Flowchart of Dataset preparation

B. System Architecture and Design

The system architecture and design have been described through step-2, step-3, step-4, and step-5 (according to Figure 1). These four steps are described in the following subsections:

a) ML Classification Algorithms: Machine learning is a branch of artificial intelligence that employs statistics, probabilities, absolute conditionality, boolean logic, and unconventional optimization strategies to learn and classify patterns through predictive models [24]. ML has both supervised and unsupervised models for classification and regression problems. Thus, we used two most commonly used supervised ML classifiers or algorithms: Random Forest (RF), and Decision Tree (DT).

1) *Random Forest Classifier*
Random Forest (RF) is a popular and most effective supervised machine learning classifier in classification and regression problems. The prime objective of RF is to reduce classification errors. It performs by building a decision tree by taking samples randomly. To classify the output, it takes the majority of the voted result. It aggregates the decisions by taking the average results of all the trees to improve the predictive accuracy and control the over-fitting problem. The advantage of the random forest classifier is to handle the data set containing continuous variables for regression and categorical variables in classification [25]. However, it provides the best performance for categorical variables in classification problems.

2) *Decision Tree Classifier*
Decision Tree (DT) is a supervised machine-learning algorithm for classification problems. It is the most popular and widely used ML algorithm. The prime goal of this algorithm is to predict the output value considering the target value and represent the solved problem in the way of tree representation called a decision tree with a leaf node or decision node [26]. A decision tree has internal and external nodes where the internal node takes part in the decision-making to make decision. In the decision tree, leaf nodes represent the class label, and the internal nodes represent the attributes. The main objective of using the Decision Tree in this work is to predict the target class instances using the decision rule learned from the prior data. In the build decision tree, root nodes classify the instances with different features where root nodes have multiple branches and the leaf nodes represent the classification result. The Decision tree chooses a node according to the highest information gain among all the attributes. The best way to information gain is to calculate entropy. Entropy is the quantified measurement of the amount of uncertainty of random instances, as shown in equation 1, where x= random instance, xi = possible outcomes, and P(xi) = probability of possible outcomes. By the value of entropy for any random instances, we can calculate the information gain through equation 2.

$$Entropy, H(X) = -\sum_{i=1}^{n} P(x|i) log P(x|i) \quad (1)$$

$$InformationGain = 1 - Entropy \quad (2)$$

b) Hyperparameter Tunning: It is noteworthy that there are several machine learning classifiers, and every ML classifier requires different constraints, weights, or learning rates to generalize the data patterns. Failure of the appropriate parameter selection might lead to differences in the final result. For example, every iteration of the classifier resulted in a different accuracy. Therefore, to optimally solve the machine learning problem and improve classification accuracy, it is crucial to appropriate parameter selections. The best way to appropriate parameter setting is hyperparameter tuning or hyperparameter optimization. Hyperparameter tuning is the process of optimal parameter setting process for a machine learning model. It is a crucial task to implement any ML model as it directly optimizes the performance of ML classification. It allows for defining all possible parameters for testing all the combinations to maximize the classification results [27]. This optimization process also allows the use of Cross-Validation (CV) to estimate the generalization performance. A crucial aspect that needs to be mentioned is that every ML classifier has different default parameter requirements. Following the default parameter, it allows the setting of the optimization parameter. Several hyperparameter optimization methods are available to select the best parameter. Herein we have used a grid search approach to evaluate classification accuracy with different combinations of parameters using a 5-fold CV. Grid search is a popular method for parameter fitting [28] that is implemented with the Scikit-learn library. For details on how parameters influence the decision of machine learning models, we suggested the literature proposed by Wu et al. [29]. In this study, we employed Grid search as its most traditional hyperparameter optimization technique, also known as parameter sweep. It is the most effective and time-constrained procedure that takes a longer time than other optimization techniques and returns better optimization results. To limit the grid search complexity, seven relevant parameters for RF and DT have been selected. The optimal values are shown in Table 3, and Table 4 which list the selected parameters of RF, and DT with definition, default values, grid values, and the optimal values for the optimization process.

TABLE III
SELECTED HYPERPARAMETER OF RF

Random Forest Parameters				
Parameter	**Definition**	**Default**	**Grid Values**	**Optimal**
criterion	The quality of a split	Gini	Gini, entropy	Gini
max_depth	Depth of a tree	None	None	None
max_features	Maximum feature of a tree	auto	auto, sqrt, log2	auto
min_samples_leaf	Minimum sample in a leaf node of a tree	1	1, 5, 8, 10, 15, 20	1
min_samples_split	Minimum samples to be split	2	2, 5, 10, 15, 20, 25	2
n_estimators	The number of estimators	100	100, 110, 120, 130, 140, 150	130
random_state	The random state of a node	None	None	None

TABLE IV
SELECTED HYPERPARAMETER OF DT

Decision Tree Parameters				
Parameter	**Definition**	**Default**	**Grid Values**	**Optimal**
criterion	The quality of a split	Gini	Gini, entropy	Gini
max_depth	Depth of a tree	None	None	None
max_features	Maximum feature of a tree	None	None	None
max_leaf_nodes	The maximum leaf node of a tree	None	None	None
min_samples_leaf	Minimum sample in a leaf node of a tree	1	1,2,3,4,8,12, 18,20	2
min_samples_split	Minimum samples to be split	2	2,3,4,5,8,10, 15	2
random_state	The random state of a node	None	None	None

c) Cross-Validation (CV): The most effective way to validate the performance of an ML model is to train a model with available data and test its classification performance using a newly separated dataset. Another popular technique is the Train-Test Split. It is the process of data splitting before model development and using the separated data for performance validation. However, these processes require a substantial amount of data for validation. Generally, CV is used to evaluate the performance of learning algorithms or models by partitioning data into a training set for pattern learning and a testing set for model evaluation [26]. The prime idea is to split the dataset into training and test sets according to the user-defined number of partitions such as K-fold. First, the dataset is divided into k folds where k-1 fold is used for training and the remaining fold for testing. In our experiment, we split the entire dataset into 70:30, where 70% of the data was used for training and 30% for testing purposes. We trained our model on the training set considering a five-fold partitioning setting and evaluated the model through a testing set for performance measurement. For hyperparameter tuning, we also applied five-fold partitioning on each fold of the training set.

d) Performance Evaluation Metrics: The evaluation metrics for measuring the classifier performance are derived from the binary confusion matrix, as represented in Figure 3. We employed True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN) counts for calculating the classification accuracy, precision, recall, and f1 score of the classifier. TP, FP, FN, and TN have been explained in following through equations 3-7.

True Positive (TP): Number of predicted instances as positive which are originally positive.

False Positive (FP): Number of predicted instances as positive which are originally negative.

True Negative (TN): Number of predicted instances as negative which are originally negative.

False Negative (FN): Number of predicted instances as negative which are originally positive.

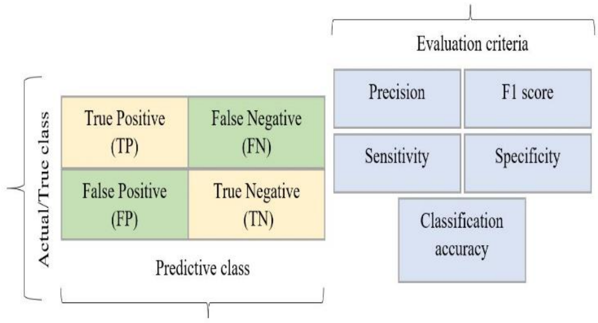


Fig. 3. Confusion Matrix with several evaluation metrics

Classification accuracy: Classification accuracy is the proportion of the number of correctly classified samples (Equation 3).

$$Classification\ Accuracy = (TP + TN)/(TP + TN + FP + FN) \quad (3)$$

Precision: Precision is the proportion of the samples that are actually true (Equation 4).

$$Precision = TP/(TP + FP) \quad (4)$$

Sensitivity: Sensitivity is the proportion of total correctly predicted samples by the learning algorithm (Equation 5).

$$Sensitivity = TP/(TP + FN) \quad (5)$$

Specificity: Specificity is the proportion of the correctly predicted negative sample with all negative samples (Equation 6).

$$Specificity = TN/(TN + FP) \quad (6)$$

F1 score: F1 score is the harmonic mean of precision and recall. To achieve the best performance, F1 score should be one, and for the lowest performance, it's usually zero (Equation 7).

$$F1\ Score = 2 * (Precision * Sensitivity)/(Precision + Sensitivity) \quad (7)$$

IV. Experimental Analysis

This section discusses the experimental setting, evaluation of classifier performance, and evaluation of website performance. The primary aim of this section is to investigate the effects of five datasets with different ratios on selected ML classifiers to verify their performance. This task has accomplished thorough experiments on multiple datasets using selected ML classifiers. Then ML classifier is applied to identify the effects of the dataset and derive the performance of the applied classifier considering various evaluation metrics.

A. Experimental Setting

The experiments with different datasets were performed considering Python programming language in a jupyter notebook environment. The experiments were run on a computer with an integrated 2.5 GHz processor and 8 GB RAM. We employed two ML classifiers: Random Forest and Decision Tree for their efficiency and reliability [30]. Every experiment was performed by taking datasets in CSV format as input to the ML classifier and splitting the dataset into a 70:30 ratio.

B. Performance Evaluation and Accessibility Score Computation

The confusion matrix of experimented datasets is tabulated in Table 5 for two selected classifiers: Random Forest (RF), and Decision Tree (DT), respectively. Table 5 also elucidates different measurements for classification result assessment, such as precision, sensitivity, specificity, F1 score, and overall accuracy to illustrate the effectiveness of the selected machine learning classifier. This table illustrates that the Random Forest (RF) classifier performs well for all five tested datasets.

TABLE V
CLASSIFICATION RESULTS OF FIVE DATASETS USING RF AND DT

Random Forest (tested data)									
		Accessible (0)	**Not Accessible (1)**	**Partially Accessible (2)**	**Classification Metrics**				
Datasets	**Actual class**	**Predictive class**			**Precision**	**Sensitivity**	**Specificity**	**F1 score**	**Accuracy**
Dataset-1	Accessible (0)	**3**	0	0	1.0	0.95	0.99	1.0	0.97
	Not Accessible (1)	0	**9**	0					
	Partially Accessible (2)	0	2	**9**					
Dataset-2	Accessible (0)	**3**	0	0	1. 0	0.86	0.99	1.0	0.99
	Not Accessible (1)	0	**8**	0					
	Partially Accessible (2)	0	0	**12**					
Dataset-3	Accessible (0)	**5**	0	0	1. 0	0.86	0.99	1.0	0.97
	Not Accessible (1)	0	**4**	0					
	Partially Accessible (2)	0	2	**12**					
Dataset-4	Accessible (0)	**1**	0	0	1.0	0.94	0.99	0.97	0.99
	Not Accessible (1)	0	**7**	0					
	Partially Accessible (2)	0	0	**15**					
Dataset-5	Accessible (0)	**4**	0	0	1.0	1.0	1.0	1.0	1.0
	Not Accessible (1)	0	**9**	0					
	Partially Accessible (2)	0	0	**10**					
Decision Tree (tested data)									
Dataset-1	Accessible (0)	**2**	0	0	0. 97	0.97	0.97	0. 96	0.95
	Not Accessible (1)	0	**8**	2					
	Partially Accessible (2)	0	1	**10**					
Dataset-2	Accessible (0)	**4**	0	0	0.98	0.95	0.93	0.95	0.94
	Not Accessible (1)	0	7	3					
	Partially Accessible (2)	0	0	**9**					
Dataset-3	Accessible (0)	**5**	0	0	1.0	1.0	1.0	1.0	1.0
	Not Accessible (1)	0	**5**	0					
	Partially Accessible (2)	0	0	**13**					
Dataset-4	Accessible (0)	**2**	0	0	0.96	0.94	0.98	0.98	0.94
	Not Accessible (1)	0	**12**	0					
	Partially Accessible (2)	0	1	**8**					
Dataset-5	Accessible (0)	**4**	0	0	0.91	0.92	0.94	0.94	0.92
	Not Accessible (1)	0	**6**	3					
	Partially Accessible (2)	0	0	**10**					

To compute the accessibility score, we quantify the score of each class (0, 1, 2) based on the number of samples of predicted data (Table 5) as shown in Equation 8. We set the severity score based on the importance of three classes as shown in Equation 9. The computed score of each class has been scaled down by multiplying their severity level as shown in Equation 10 and computing the final score through Equation 11.

$$Accessible\ (\alpha) = TP_0,\ Partially\ Accessible\ (\beta) = TP_1,\ Not\ Accessible\ (\Upsilon) = TP_2 \quad (8)$$

$$\epsilon_\alpha = 0.2,\ \epsilon_\beta = 0.1, \epsilon_\Upsilon = 0.01 \quad (9)$$

$$Accessible = [\alpha * \epsilon_\alpha], Partially\ Accessible = [\beta * \epsilon_\beta],\ Not\ Accessible = [\Upsilon * \epsilon_\Upsilon] \quad (10)$$

$$Total\ accessibility\ score = \frac{\{(Accessible + Partially\ Accessible) - Not\ Accessible\}}{n} * 100 \quad (11)$$

The computed accessibility score for each classifier with their average score and standard deviation (SD) is shown in Table 6. According to the computed accessibility score, it illustrates that dataset 3 has a higher accessibility score than other experimented datasets. Besides, dataset 4 experienced with the lowest accessibility score. In the context of accessibility, the higher the accessibility score the lower the accessibility barrier, and the lower the accessibility score the higher the accessibility barrier. Also, the standard deviation represents the difference between the computed results of two implemented classifiers. The highest difference (in terms of SD) among two classifiers results was observed for dataset 4 and dataset 5.

TABLE VI
ACCESSIBILITY SCORE COUNT OF TESTED DATASET

Dataset	RF	DT	Avg. Accessibility Score	SD
Dataset-1	47%	44%	45.5%	2.12
Dataset-2	57%	53%	55%	2.23
Dataset-3	72%	75%	73.5%	2.12
Dataset-4	25%	36%	30.5%	7.77
Dataset-5	53%	58%	55.5%	3.53

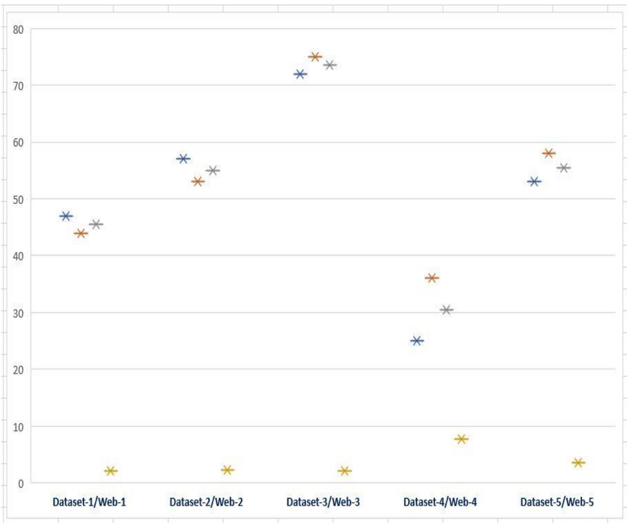


Fig. 4. Graphical representation of the computed accessibility score of five selected university websites.

Figure 4 shows the graphical representation of the computed accessibility score (for both RF and DT), their average accessibility score, and SD values for each dataset where we reference the results of each dataset with their associated dataset/website. It shows that the website of the Budapest University of Technology and Economics (Dataset 4/Web-4) has the lowest accessibility score than other selected university websites. In contrast, the website of Eotvos Lorand University (Dataset 3/Web-3) has the highest accessibility score. From these computed scores, it is noteworthy that none of the selected university websites were found completely accessible or barrier-free for people with several disabilities in accordance with our selected ten features.

V. DISCUSSION

This work presents an extensive study of university website accessibility score computation in Hungary. Generally, university websites are the prime source of information and services for native and international students and stakeholders. Thus, there’s an emerging need to identify the accessibility status of university websites. Besides, there is no previous research work conducted in Hungary to evaluate the university websites or academic websites of this country. Addressing these gaps, we proposed a web accessibility evaluation approach considering machine learning algorithms to compute the accessibility score of the selected university web pages of Hungary.

Generally, web content accessibility guidelines are widely accepted standards but few issues or aspects associated with people with disabilities are not considered in this standard. For example, if a website does not provide a manual text size adjustment option or manual color adjustment option then the majority of the people with vision disability or color disabilities will face difficulty in navigating the content. Besides, a few websites require user information for browsing some specific content, and few require successful completion of CAPTCHA which is very difficult for people with cognitive disabilities. Some other issues with the use of less-productive/not useful internal/external links, images, and video and audio content also hampered the accessibility aspects for people with special needs. Unfortunately, most of the automated accessibility testing tools also do not consider these accessibility aspects as these issues are very complicated to incorporate into an automated manner [31]. Therefore, accessibility checking considering these aspects might be useful for revealing the true insights of website accessibility.

With this aim, we computed the accessibility score by incorporating the Random Forest (RF) and Decision Tree (DT) techniques and calculating their average score with individual SD calculations. Our experiment result shows that the classification performance of the Random Forest classifier is more significant than the Decision Tree classifier. The average accessibility score shows that Eotvos Lorand University has higher accessibility features (according to the selected features in this research work) than other university websites. However, the computed score of other selected university websites was very poor which represents that most of the selected university web pages are not accessible to people with disabilities in terms of the selected aspects/features. To improve the accessibility in accordance with the selected aspects, tested university web pages need to improve their quality to ensure the complete accessibility objective. In addition, concerning the performance of machine learning classifiers or models, it is interesting to address that machine learning classifiers or models are significant in the evaluation of the accessibility of university websites. However, throughout the experimentation, we have some limitations associated with single-page validation and a small dataset. Therefore, a further investigation is required focusing on the current limitation that will be considered in future work.

VI. Conclusion

The prime objective of this study is to present the accessibility insights of higher institute websites (university websites) by implementing machine learning methods. University websites are an emerging platform to distribute information to students and associated authorities. However, providing an accessible online or accessible website is a relatively challenging task for the web designer and developer. Literature shows that few studies focused on accessibility issues, and there is almost no research work conducted considering the higher institute websites of Hungary. Therefore, we proposed a machine learning approach for computing the accessibility barrier score of the selected university website. We evaluated the result of machine learning methods through several metrics obtained from confusion matrix and classification reports. The computed result predicts that the selected university websites are not free from accessibility issues that reduce their effectiveness. The future work is limited to three focused groups: (i) this work is limited to incorporating two ML models which will be further extended and incorporated with other ML models focusing on the current limitations, (ii) to validate the result, we will incorporate user/ expert testing, and (iii) the entire university websites will be considered for experimenting instead of considering a single webpage/homepage.

References

[1] Allison, R., Hayes, C., McNulty, C. A., & Young, V. (2019). A comprehensive framework to evaluate websites: literature review and development of GoodWeb. *JMIR formative research*, *3*(4), p. e14372.

[2] Longstreet, P., Brooks, S., Featherman, M., & Loiacono, E. (2021). Evaluating website quality: which decision criteria do consumers use to evaluate website quality?. *Information Technology & People*.

[3] Vargas, A. M., Pedraza-Jiménez, R., & Bonilla, L. C. (2020). Website quality: An analysis of scientific production. *El profesional de la información*, *29*(5), p. 22.

[4] Acosta-Vargas, P., Acosta, T., & Lujan-Mora, S. (2018). Challenges to assess accessibility in higher education websites: A comparative study of Latin America universities. *Ieee Access*, *6*, pp. 36 500–36 508.

[5] Dominic, P. D. D., Jati, H., Sellappan, P., & Nee, G. K. (2011). A comparison of Asian e-government websites quality: using a non-parametric test. *International Journal of Business Information Systems, 7*(2), pp. 220–246.

[6] Hackett, S., Parmanto, B., & Zeng, X. (2005). A retrospective look at website accessibility over time. *Behaviour & Information Technology*, *24*(6), pp. 407–417.

[7] Ringlaben, R., Bray, M., & Packard, A. (2014). Accessibility of American university special education departments' web sites. *Universal Access in the Information Society, 13*(2), pp. 249–254.

[8] Devi, K., & Sharma, A. (2016). Framework for evaluation of academic website. International Journal of Computer Techniques, 3(2), pp. 234-239. [9] Montazer, G. A. (2018). University website quality improvement using intuitionistic fuzzy preference ranking model. *Quarterly of Iranian Distance Education Journal, 1*(2), pp. 9–30.

[10] Kaur, S., Kaur, K., & Kaur, P. (2016). An empirical performance evaluation of universities website. *International Journal of Computer Applications, 146*(15), pp. 10–16.

[11] Yerlikaya, Z., & Onay Durdu, P. (2017, July). Evaluation of accessibility of university websites: A case from turkey. In *International Conference on Human-Computer Interaction* (pp. 663–668). Springer, Cham.

[12] Ara, J., & Sik-Lanyi, C. (2023, May). AccGuideLiner: Towards a Modelling Approach of Web Accessibility Requirements following WCAG 2.2. In *2023 IEEE International Conference on Smart Information Systems and Technologies (SIST)* (pp. 10–15). IEEE.

[13] Chopra, G., Madan, P., Jaisingh, P., & Bhaskar, P. (2019). Effectiveness of e-learning portal from students' perspective: A structural equation model (SEM) approach. *Interactive Technology and Smart Education*.

[14] Mittal, H., Sharma, M., & Mittal, J. P. (2012, January). Analysis and modelling of websites quality using fuzzy technique. In *2012 Second International Conference on Advanced Computing & Communication Technologies* (pp. 10–15). IEEE.

[15] Malhotra, R., & Sharma, A. (2013, August). A neuro-fuzzy classifier for website quality prediction. In *2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 1274–1279). IEEE.

[16] Rashida, M., Islam, K., Kayes, A. S. M., Hammoudeh, M., Arefin, M. S., & Habib, M. A. (2021). Towards developing a framework to analyze the qualities of the university websites. *Computers, 10*(5), pp. 57.

[17] Olaleye, S. A., Sanusi, I. T., Ukpabi, D., & Okunoye, A. (2018). Evaluation of Nigeria universities websites quality: A comparative analysis.

[18] Alahmadi, T. (2017, April). A multi-method evaluation of university website accessibility: Foregrounding user-centred design, mining source code and using a quantitative metric. *In Proceedings of the 14th International Web for All Conference* (pp. 1–2).

[19] Hassouna, M. S., Sahari, N., & Ismail, A. (2017). University website accessibility for totally blind users. *Journal of Information and Communication Technology, 16*(1), pp. 63–80.

[20] Verkijika, S. F., & De Wet, L. (2020). Accessibility of South African university websites. *Universal Access in the Information Society, 19*(1), pp. 201–210.

[21] AlMeraj, Z., Boujarwah, F., Alhuwail, D., & Qadri, R. (2021). Evaluating the accessibility of higher education institution websites in the State of Kuwait: empirical evidence. *Universal Access in the Information Society, 20*(1), pp. 121–138.

[22] Dhiman, P. (2014, September). Empirical validation of website quality using statistical and machine learning methods. In *2014 5th International Conference-Confluence The Next Generation Information Technology Summit (Confluence)* (pp. 286–291). IEEE.

[23] Hungary Ranking Web of Universities-Webometrics: https://www.webometrics.info/en/europe/hungary

[24] Raschka, S. (2018). Model evaluation, model selection, and algorithm selection in machine learning. *arXiv* preprint *ArXiv*:1811.12808.

[25] Izquierdo-Verdiguier, E., & Zurita-Milla, R. (2020). An evaluation of Guided Regularized Random Forest for classification and regression tasks in remote sensing. *International Journal of Applied Earth Observation and Geoinformation, 88*, p. 102051.

[26] Sisodia, D., & Sisodia, D. S. (2018). Prediction of diabetes using classification algorithms. *Procedia computer science, 132*, pp. 1578–1585.

[27] Vabalas, A., Gowen, E., Poliakoff, E., & Casson, A. J. (2019). Machine learning algorithm validation with a limited sample size. *PloS one, 14*(11), p. e0224365.

[28] Yang, L., & Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing, 415*, pp. 295–316.

[29] Matošević, G., Dobša, J., & Mladenić, D. (2021). Using machine learning for web page classification in search engine optimization. *Future Internet, 13*(1), p. 9.

[30] Wu, J., Chen, X. Y., Zhang, H., Xiong, L. D., Lei, H., & Deng, S. H. (2019). Hyperparameter optimization for machine learning models based on Bayesian optimization. *Journal of Electronic Science and Technology, 17*(1), pp. 26–40.

[31] Ara, J., & Sik-Lanyi, C. (2022). Investigation of COVID-19 Vaccine Information Websites across Europe and Asia Using Automated Accessibility Protocols. *International Journal of Environmental Research and Public Health, 19(5)*, p. 2867.

Jinat Ara graduated from Southeast University, Bangladesh in 2018 with a bachelor's degree in Computer Science and Engineering (CSE). Following that she obtained her master's degree in 2020 from Jahangirnagar University, Bangladesh from the Department of Computer Science (CS). At the same time, she worked as a Software Engineer at Odyssey Apps, Dhaka, Bangladesh. Currently, she is pursuing her Ph.D. degree at the Faculty of Information Technology, University of Pannonia, Hungary. Her research area for her Ph.D. studies on the area of "accessibility investigation in web platforms or environments". Additionally, she has a keen interest in conducting cutting-edge research and currently, she is carrying around six years of research experience in several areas of computer science. Her area of research interest is 'Machine Learning (ML)', 'Virtual and Augmented Reality', 'Assistive Technologies', 'Web Technologies', 'Artificial Intelligent (AI)', 'Natural Language Processing (NLP)' and 'Automated Vehicles (AV)'.

Cecilia Sik-Lanyi is a Professor at University of Pannonia, Veszprem, Hungary. Sik-Lányi studied Mathematics and Computer Science. She obtained the degree of Dr. Univ. in Physical-chemistry (1993), and Ph.D. in Computer Science (2000), and became a Doctor of Hungarian Academy Sciences (D.Sc.) in 2022. Her research area is virtual reality, serious games, and user interface design, focusing on color science and accessibility. She received several awards, the most important ones are: "Diamond-Award from the Association for the Advancement of Assistive Technology in Europe" in 2015, and the "King Salman Award for Disability Research" in 2018. She has supervised more than 220 BSc and MSc theses since 1997. She has published more than 500 referred articles and conference papers and worked as a guest editor for many renowned journals and books.

Optimizing Text Clustering Efficiency through Flexible Latent Dirichlet Allocation Method: Exploring the Impact of Data Features and Threshold Modification

Erzsébet Tóth, and Zoltan Gal

***Abstract*—A parallel corpus comprising Croatian EU legislative documents automatically translated into English spans 28 years and is enriched with metadata, including creation year and hierarchical classifier tags denoting descriptors, document types, and fields. However, nearly two-thirds of the approximately 1.5 thousand texts lack complete metadata, necessitating labor- intensive manual efforts that pose challenges for human administration. This incompleteness issue can be observed in the case of official legal sites functioning as regular service provisioning databases. In response, this paper introduces an artificial cognitive and multilabel classification approach to expedite the tagging process with only a fraction of the manual effort. Leveraging the Latent Dirichlet Allocation (LDA) algorithm, our method assigns field values or tags to incompletely labeled documents. We implement a Flexible LDA variant, incorporating the influence of topics close to the most probable topic, regulated by a relative probability threshold (RPT). We evaluate the LDA prediction's dependence on document prefiltering and RPT values. Furthermore, we investigate the dependence of quantitative linguistic properties on the type and speciality of pre-processing tasks. Our algorithm, built on error-correcting optimizing codes, successfully predicts a mixture of topic probabilities for these legal texts. This prediction is achieved by calculating the Hamming distance of binary feature vectors created using the legal fields of the EUROVOC multilingual thesaurus.**

***Index Terms*—Multilabel classification, Legal text clustering, Latent Dirichlet Allocation, Supervised learning, Artificial Intelligence, Natural Language Processing, Quantitative linguistics.**

I. Introduction

Our research paper focuses on the cluster analysis of the Croatian and English parallel legal corpus included in the MARCELL (Multilingual Resources for CEF.AT in the Legal Domain) corpus. Previously the MARCELL corpus and its related resources were created in the MARCELL CEF (Connecting Europe Facility) Telecom Action. The CEF Telecom project Multilingual Resources for CEF.AT in the Legal Domain (MARCELL) intends to improve the eTranslation system implemented by the European Commission by providing seven large-scale corpora comprising national legislative documents effective in Poland, Bulgaria, Croatia, Romania, Hungary, Slovakia, and Slovenia [1].

We think that our paper has a close connection with the cognitive infocommunications (CogInfoCom) interdisciplinary research field because one of its main goals is to support the effective interaction between computers and humans and extend human cognitive capabilities with the help of infocommunications devices such an example can be a high-level artificial cognitive capability of a neural network used for text clustering. In addition, CogInfoCom's objective is to provide a systematic view of the co-evolution of human cognitive processes and infocommunication devices [23-25].

Recently there has been a growing interest in quantitative linguistic laws [2] and artificial intelligence used in text clustering. Rijsbergen emphasizes Luhn's work in automatic text analysis which assumed that word frequencies could be used for extracting words and sentences to articulate the content of a document. Let us consider f the word frequency in a particular position of the text and r the rank order of the words (i.e. the order of their frequency of occurrence), then a plot f versus r yields a hyperbolic curve. Besides this, it is a curve illustrating Zipf's law (studied intensively by Stephanie Evert in Nürnberg [21-22]) which declares that the product of word intensity values and their rank order is around constant. Luhn used this law as a null hypothesis to determine the upper and lower cut-offs of the rank order of words. He interpreted the significance of the words as their ability to express the topic of the text. With the help of these arbitrarily specified cut-offs, he omitted insignificant words such as rare and common words from the rank order of the items and in this way, he could identify those important words which describe the content or topic of the text [3].

Highlights of the paper are the following: a) Supervised learning based on the Latent Dirichlet Allocation (LDA) method is applied to 1119 legal EU English texts to classify 392 legal texts without having field attributes; b) The used artificial cognitive and multilabel classification was able to assign fields to the 392 EU legal texts successfully; c) It is proved that the LDA is less sensitive to the cleaning state of the analysed texts. The paper unfolds as follows: Section two provides an overview of related work on the Croatian-English legal corpus. In section three, we detail the applied methodology, delve into data

E. Tóth is with Faculty of Informatics, University of Debrecen, Debrecen, Hungary (e-mail: toth.erzsebet@inf.unideb.hu).

Z. Gal is with Faculty of Informatics, University of Debrecen, Debrecen, Hungary (e-mail: gal.zoltan@inf.unideb.hu)

DOI: 10.36244/ICJ.2024.5.7

processing elements, and engage in a discussion of results using quantitative linguistics approaches. Section four encapsulates conclusions drawn from our findings and outlines potential avenues for the continuation of this research work.

II. Related Works

Numerous scientific papers address challenges in automatically classifying texts based on the inherent features of document context. This section aims to bridge theoretical considerations with practical applications, offering key insights into context interpretation within the legal domain. We address current issues by presenting a non-exhaustive list of notable projects and analysis mechanisms in the legal field.

During the period from 2018 to 2020, the MARCELL project pursued the goal of providing fresh monolingual training material for CEF.AT Neural Machine Translation Services facilitated by the European Commission. This endeavor led to the organization of the introductory workshop for the CEF-project EU Council Presidency Translator in Zagreb in 2020. The workshop discussed the initial outcomes of the Machine Translation (MT) tool developed for English-Croatian and Croatian-English directions by the University of Zagreb, Faculty of Humanities and Social Sciences, Zagreb (Croatia), and Tilde, Riga (Latvia). The participants included professionals in translation and communication from diverse Croatian industry sectors and public authorities. The event also attracted Croatian translators from various EU bodies. Marko Tadić, a member of the MARCELL project and contributor to the development of the MT system, highlighted MARCELL as a significant source of freely available language resources for training similar MT systems used by CEF.AT users and translators [4].

The EU Council Presidency Translator toolkit, an MT service developed for the 2020 Croatia's Presidency of the Council of the European Union, operates as a multilingual communication tool, facilitating instant translation of texts, documents, and websites between Croatian and English. Utilizing neural networks enhanced with artificial intelligence and machine learning, this toolkit is accessible online or through SDL Trados Studio with a plug-in, powered by the CEF eTranslation platform [5][6].

In the realm of multilabel classification, a noteworthy contribution is found in the paper [18]. The authors propose a novel method enhancing multilabel classifier performance by considering label correlations. The paper provides a comprehensive description of the classifier chains method, comparing it with various multilabel classification methods across diverse datasets. Experimental results demonstrate the superiority of the classifier chains method in multiple metrics, including precision, recall, and F1-score.

Another paper introduces Label-Specific Feature Learning (LSFL), a method involving the acquisition of label-specific feature vectors capturing the characteristics of each label [19]. These feature vectors are integrated with input feature vectors, forming a new representation for each instance, subsequently used to train a multilabel classifier. LSFL, based on regularized matrix factorization, optimizes the input feature matrix and label-specific feature matrix, preventing overfitting and encouraging sparsity. The LSFL method, as reported, outperforms several state-of-the-art multilabel classification methods across various evaluation metrics.

A widely-cited paper reviews diverse approaches to multilabel learning, encompassing problem transformation methods, algorithm adaptation methods, and ensemble methods [20]. The authors discuss evaluation measures, including precision, recall, F1-score, and Hamming loss, presenting experimental results on benchmark datasets to compare algorithm performance based on these measures.

III. Applied Methodology, Data Processing and Discussion

In the next subsections, we describe the input data set applied in the multilabel classification process [7][8]. Topic discovery with the Latent Dirichlet Allocation method will be detailed in [9][10][11]. Interpretation of the results will be based on quantitative linguistic approaches [12][13].

A. Description of the Data

Upon Croatia's accession to the EU in 2013, the country initiated the translation of national legislative documents between Croatian and English, resulting in the creation of the Croatian-English Parallel Corpus of Croatian National Legislation spanning texts from 1990 to 2019, totaling approximately 1,800 documents. Notably, earlier Croatian legal documents were monolingual, and the English translations commenced in PDF format in the late 1990s.

However, the extraction of text from various PDF files posed challenges, impacting the quality of the automatic extraction process. To address this, the researchers employed sentence splitting and alignment using LF-aligner [14], an open-source tool utilizing HunAlign in the background [15]. To ensure high-quality alignment, proofreading was manually conducted for all 1,816 documents, resulting in accurately aligned TMX files. This meticulously curated parallel corpus serves as a valuable resource for the noiseless training of Neural Machine Translation (NMT) systems. The corpus encompasses 396,984 token units, with 14.4 million and 17.7 million tokens in Croatian and English, respectively.

In general, the 1816 source documents collected from 1991 to 2018 (28 years) have a header with a specific structure and a body with Croatian and corresponding English split sentences. The header consisted of the document type (noted T attribute), year of creation (noted Y attribute), EUROVOC descriptors (noted D attribute) and field (noted F attribute) elements. The number of unique entities of the document type, year, descriptor and field elements is $|T| = 11, |Y| = 28, |D| = 1393$ and |F| = 22, where $|X|$ represents the cardinality of the document subset with attribute $X \in \{T, Y, D, F\}$. In this data set, there are 1585 documents having type and year in the header (noted TY attributes). A subset of 1511 documents has a type, year in the header and sentences in the body (noted $TY{:}S$ attributes). Only 1119 documents have type, year, field in the header and sentences in the body (noted $TYF{:}S$ attributes). In this way, the number of documents without field is $1511 - 1119 = 392$, called a set of unlabelled documents [16]. In our analysis, the document ID range is $1 \ldots 1511$ in chronological order.

TABLE I.
LIST OF DOCUMENT TYPES (T).

ID	Type	ID	Type
1	Constitution	7	Order
2	Constitutional act	8	Ordinance
3	Decision	9	Other
4	Instructions	10	Regulation
5	Law	11	Standing orders/rules of procedure
6	Legal code		

Table 1 provides a list of distinct document types, aligning with the CELEX identification system of the EU. The CELEX database stands out for its expansive coverage and advanced search functionalities, making it an invaluable resource for legal research and analysis. Users can conduct searches based on document type, date, subject matter, and keywords, as well as utilize specific CELEX numbers assigned to each document.

Table 2 summarizes the unique fields within the legal corpus, with each field serving as a label derived from the EUROVOC multilingual, hierarchical thesaurus developed by the Publications Office of the EU. The collection comprises a total of 22 fields, representing various subject areas. This diversity enables users to navigate and search for terms and concepts pertinent to specific domains within different official information systems.

TABLE II.
LIST OF DOCUMENT FIELDS (F).

ID	Field
1	Agriculture, forestry, fishery
2	Communication and information technology
3	Construction and city planning
4	Culture and cultural property
5	Defence, internal affairs and national security
6	Economy, trade and commerce
7	Education and sports
8	Energy production
9	Environment and natural heritage
10	Finance, budget and monetary affairs
11	Health care
12	Industry and technology
13	Information, media, documentation, statistic
14	International relations and cooperation
15	Labour, employment and pension scheme
16	Law and the judiciary
17	Politics and public authority
18	Science and research
19	Social activities and human rights
20	Social care
21	Tourism and tourist activities
22	Traffic and traffic infrastructure

The distribution of the 1511 documents with type, year and sentences is represented on the left side of Fig. 1. according to the analysed years. The histogram of the number of documents vs. years is illustrated on the right side of Fig. 1.

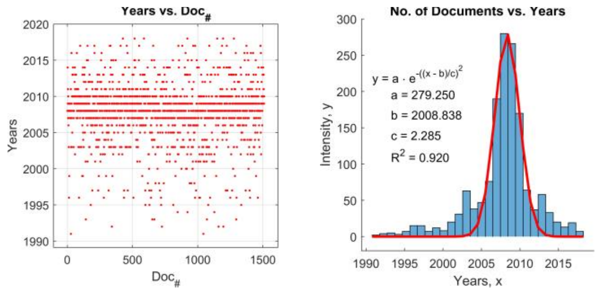


Fig. 1. Years vs. DocID (left); No. of documents vs. years (right)

The largest number of documents was created around the year 2008. The fitting curve of the intensity versus years has the following equation:

$$Intensity = a \cdot e^{-[(year - b)/c]^2} \quad (1)$$

Values of the parameter triplet (a, b, c) are $(279.25, 2008.84, 2.28)$ with 92% coefficient of determination. The independent variable takes values in the following interval: $year \in \{1991, \ldots, 2018\}$.

In Fig. 2, the left side presents an overview of the potential number of descriptors per document yearly, while the right side illustrates the distribution of descriptors per document across the years. Notably, the number of descriptors per labeled document remains consistently below a dozen for the examined years. However, when considering both labeled and unlabeled documents, the average number of descriptors is less than four. This difference arises due to a substantial portion of documents in the English-Croatian legal corpus lacking descriptors.

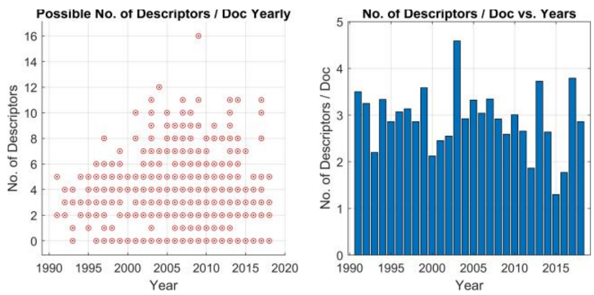


Fig. 2. Possible No. of descriptors/doc yearly (left);
No. of descriptors/doc vs years (right).

Note that no more than 16 descriptors were assigned to any of the documents belonging to the 28 years of the survey (see Fig. 3 left). Most of the documents, $\sim 25\%$ have just one descriptor and $\sim 70\%$ of them have between 3 and 7 descriptors (see Fig. 3 right).

Because the number of descriptors is very high ($|D| = 1393$), this attribute was not considered in this multilabel analysis [17]. In contrast to descriptors, the number of fields is just 22 creating the possibility to use them as a multilabel binary vector of 22 dimensions.

We define a topic of the document by the field pattern of its feature vector. The majority of documents exhibit one of two fields, but a significant portion of legal texts lacks fields altogether (refer to Fig. 4, left).

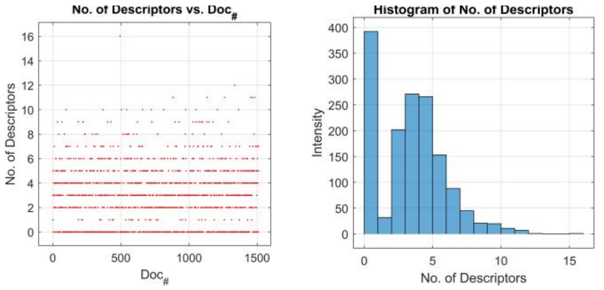


Fig. 3. No. of descriptors vs. docID (left);
Histogram of No. of descriptors (right).

Overall, the average number of fields per document is predominantly less than one during the examined years (refer to Fig. 4, right).

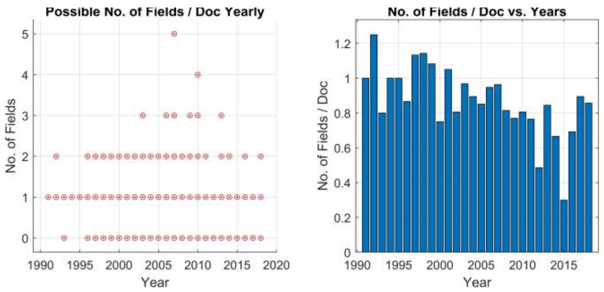


Fig. 4. Possible No. of fields/doc yearly (left);
No. of fields/doc vs. years (right).

This reflects that the EUROVOC thesaurus was not exploited sufficiently for assigning labels to these documents. A maximum of five fields are assigned to the investigated texts and the majority (~66%) of them have just one field. The number of unlabelled items is 392 (see Fig. 5).

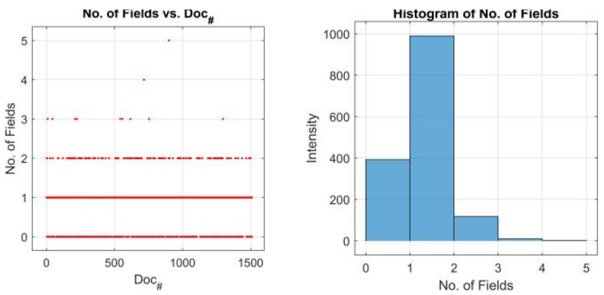


Fig. 5. No of fields vs. docID (left);
Histogram of No. of fields (right).

We observed that most of the items have around 100 sentences and the longest item has less than 5000 sentences (see Fig. 6 left). The histogram of the sentences conforms to an exponential function:

$$Intensity = a \cdot e^{-b \cdot year} \tag{2}$$

Values of the parameter triplet (a, b) are $(1316.183, -0.007)$ with 99% coefficient of determination. The independent variable takes values in the following interval: $year \in \{1991, \dots, 2018\}$ (see Fig. 6 right).

There was a high number of sentences created from 2007 to 2012 (see Fig. 7 left). The largest item has less than 500 and the shortest one has at least 30 sentences, respectively. Note that from 2007 to 2021 a higher amount of documents was created resulting in a lower number of sentences per document (see Fig. 7 right).

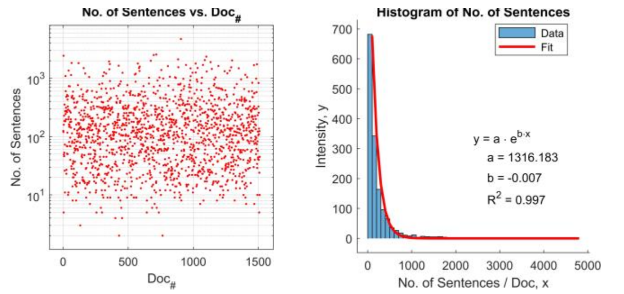


Fig. 6. No of sentences vs. docID (left);
Histogram of No. of sentences (right).

The longest documents were created in the years of 1995 and 2013.

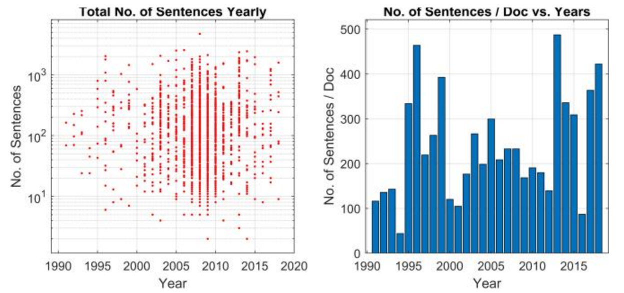


Fig. 7. Total No of sentences yearly (left);
No. of sentences/doc vs. years (right).

The histogram of the sorted unique descriptors follows an exponential equation:

$$Intensity = a \cdot e^{-b \cdot x^{c}} \tag{3}$$

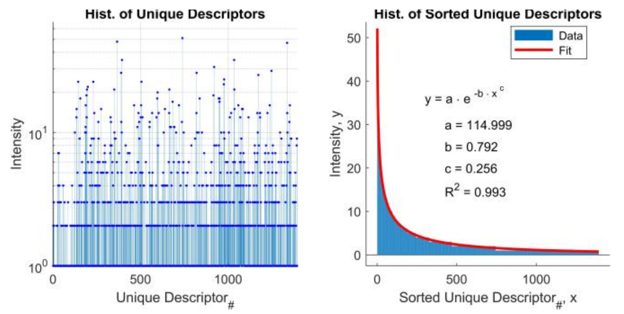


Fig. 8. Histogram of unique doc types (left);
Histogram of sorted unique descriptors (right).

Values of the parameter triplet (a, b, c) are $(114.99, 0.79, 0.26)$ with 99% coefficient of determination. The intensity of the unique descriptors covers the range of 1 and 50 and they have specific occurrences in the corpus (see Fig. 8 left). The sorted unique descriptor (independent variable), x takes values in the following interval: $x \in \{1, \dots, 1393\}$ (see Fig. 8 right).

Most of the documents conform to type 8 (Ordinance), followed by type 5 (Law) and type 3 (Decision) of Table I. (see Fig. 9 left). The first two most frequent fields of the documents listed in Table II. belong to the "10: Finance, budget and monetary affairs" and "1: Agriculture, forestry, fishery" domains having ratio of approximately 30% and 20%, respectively (see Fig. 9 right).

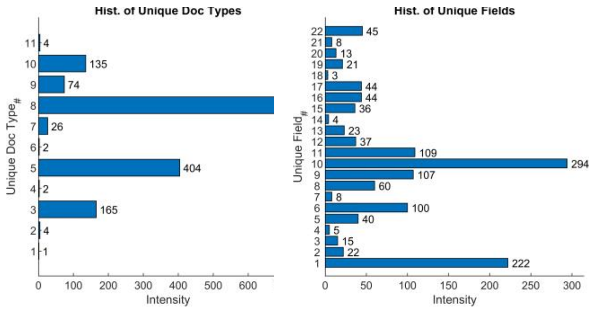


Fig. 9. Histogram of unique doc types (left);
Histogram of unique fields (right).

The smallest intensity of the fields was “18: Science and research”, “14: International relations and cooperation” and “4: Culture and cultural property”. In the next section, we highlight the main aspects of the LDA method and its application in our data set.

B. Topic Discovery with Latent Dirichlet Allocation Algorithm

LDA operates as a hierarchical Bayesian model featuring three levels. The items within the collection set constitute a blend of topic probabilities, where each topic represents an infinite amalgamation of the basic set of topic probabilities. In the context of a document, the likelihood of a topic imparts characteristics to the text corpus. The Hamming distance, a straightforward and intuitive distance metric, finds utility in various applications, including clustering, classification, and information retrieval. It serves to compare document similarity based on labels, contexts, or themes.

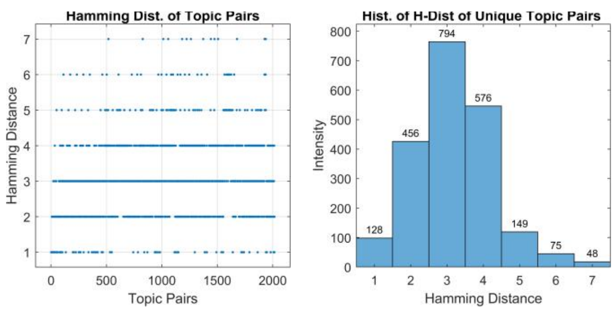


Fig. 10. Hamming distance of topic pairs (left);
Histogram of Hamming distance of unique topic pairs (right).

Our approach involved utilizing the Hamming distance between topic pairs and their corresponding histogram to assess the dissimilarity of label assignments across the texts (refer to Fig. 10). The maximum distance observed between topics is seven, with the majority of topic pairs positioned at a distance of three. The overall distribution converges towards a Gaussian bell shape (refer to Fig. 10, right).

Feature binary vectors of two sets of documents with cardinality 30 are represented in Fig. 11, respectively. The square means bit one in the corresponding field position of the vector. Document IDs increase from bottom to top.

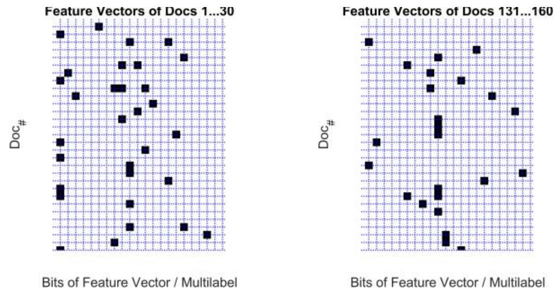


Fig. 11. Feature vectors of docs 1...30 (left);
Feature vectors of docs 131...160 (right).

The horizontal pattern refers to the binary feature vector representing the topic details of a specific document. It's worth noting that a small number of documents in both sets depicted in Fig. 11 lack a pattern, indicating the absence of fields in these texts. Our objective was to predict the missing fields in these texts by leveraging the meaning of sentences through the application of the LDA method.

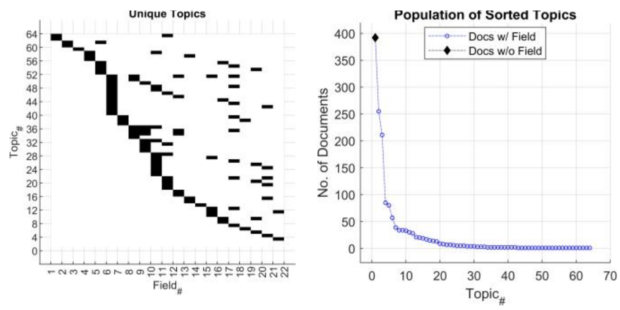


Fig. 12. Unique topics of the document corpus (left);
Population of the sorted topics (right).

Within the analysed corpus, feature vectors exhibit distinct bit patterns corresponding to a total of 64 topics (refer to Fig. 12). A binary vector featuring homogeneous 0-s is termed a pseudo-topic, serving to represent documents devoid of fields. The 392 documents associated with pseudo-topics are illustrated in the right figure using a diamond shape.

Fig. 13 (left) showcases a toplist of tokens derived from both original and cleaned texts on a log-log scale. Notably, both intensities exhibit well-approximated straight lines, suggesting power functions in the linear scale. The scatterplot of the intensity of the cleaned and original tokens shows linear dependencies in two intervals (see Fig. 13 right).

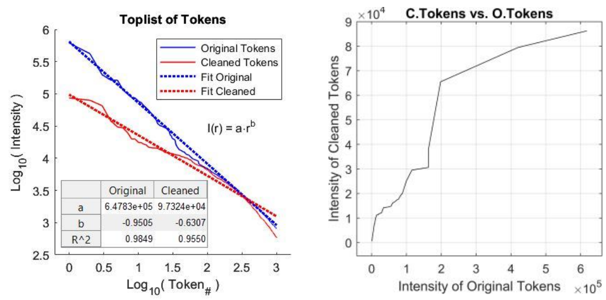


Fig. 13. Toplist of tokens (left);
Cleaned Tokens vs. Original Tokens (right)

This indicates that the cleaning ratio of the tokens is constant for the majority of the large documents. The slope of the original tokens is close to -1, corresponding to Zipf's law. However, the tokens of the cleaned texts have an exponent of -0.63 proving unconformity to Zipf's law. A possible explanation for this unconformity may be due to the cleaning task executed on the corpus. The cleaning process of the texts and usage of working objects are listed in Table III.

Tasks executed in each of the cleaning steps are as follows. 1: Convert string into tokens; 2: Add part of speech details; 3: Details of Original Tokens; 4: Reduce each token to stem; 5: Remove stopwords; 6: Remove Short (less than 3 characters) and Long Words (greater than 15 characters); 7: Details of Stem Tokens; 8: Create a bag-of-words; 9: Create a bag-of-words of cleaned docs; 10: Create TopBag of Original Tokens; 11: Create TopBag of Cleaned Tokens. Variable noTokens has a value of 1000 because we consider only the first 1000 most frequent tokens as significant.

TABLE III.
LIST OF DATA PROCESSING STEPS.

StepID	Function
1	docs = TokenizedDocument(texData)
2	docs1 = AddPartOfSpeechDetails(docs)
3	tokenDetailsOrig = TokenDetails(docs1)
4	docs2 = NormalizeWords(docs1)
5	docs3 = RemoveStopWords(docs2)
6	docsClean = RemoveShortWords(docs3, 2, 15)
7	tokenDetailsStem = TokenDetails(docsClean)
8	bag = BagOfWords(docs)
9	bagClean = BagOfWords(docsClean)
10	topBag = Topkwords(bag, noTokens)
11	topBagClean = topkwords(bagClean, noTokens)

Applying LDA for topic prediction on labeled documents yields a list of potential topics quantified by their respective probabilities of belonging. Word clouds are subsequently generated by arranging tokens in descending order of occurrence within the text. These tokens represent reduced stems of words (such as "servic," "articl," "manufactur," "measure"), with some exceptions (like "electron," "croatian," "conform," "limit").

Prediction examples of four selected texts (doc ID = 25, 26, 33, 34) in decreasing order of the belonging probability are shown in Fig. 14. right side.

We observed that some documents are unique, but others have a few dominant probabilities.

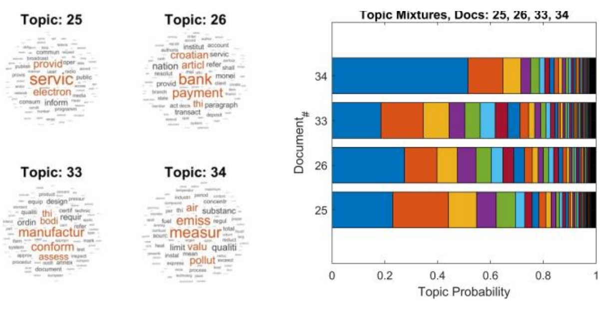


Fig. 14. Sample of topic word clouds of a text quartet (docs 25, 26, 33, 34) (left); Topic mixtures of the text quartet (right).

The latter documents weaken the goodness of the topic prediction.

C. Impact of the similarity probabilities on the LDA

To consider significant topics not just the absolute first candidate we introduced a parameter called topic Relative Probability Threshold (selection of significance threshold), $RPTh$. This parameter is used to binary classify topics into significant and non-significant groups of the normalized topic probabilities. Note that when $RPTh = 100\%$ we have the classical LDA algorithm with only one solution, which means that each of the selected four texts has only one dominant topic.

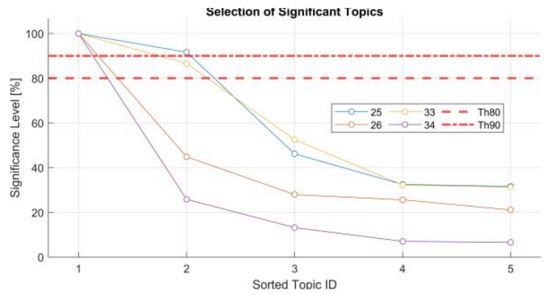


Fig. 15. Selection of significant topics of the text quartet (docs 25, 26, 33, 34).

In our case, this means selection of Topic34, Topic48, Topic2 and Topic 62 for the texts 25, 26, 33 and 34, respectively. For $RPTh \in (0,1)$ the extended LDA may have more than one topic proposal for close probability values. For $RPTh = 0$, the LDA gives all 63 combinations as a proposal of the multilabel classification problem, implying non-usability in practice.

In the case of the significance threshold of the LDA having value $RPTh80 = 80\%$, the estimation of the text quartet is shown in Table IV. Two of the documents (25, 33) have more than one estimation and others are identified by one estimation (see Fig. 15). Note that in this randomly selected text quartet, the first and second topic estimation for the same document differs strongly. The first estimation for text 33 is Topic2 ("Traffic and traffic infrastructure"). The second estimation is Topic43 ("Economy, trade and commerce", "Politics and public authority") which is quite different from Topic2. If the significance threshold value is set to $RPTh90 = 90\%$, just one text (25) has more than one multilabel estimation and text 33 gets only one dominant topic (multilabel2). This behaviour of the modified LDA mechanism proves the strong dependence of the estimation decision on the value of the significance threshold.

TABLE IV.
LIST OF SIGNIFICANT TOPICS OF THE TEXT QUARTET (DOCS 25, 26, 33, 34), THRESHOLD = 80%.

Doc ID	Significant Topic IDs
25	1st estimation : Topic34 (Field8, Field9, Field12) 2nd estimation: Topic61 (Field2, Field5)
26	1st estimation: Topic48 (Field6, Field10, Field17)
33	1st estimation: Topic2 (Field22) 2nd estimation: Topic43 (Field6, Field17)
34	1st estimation: Topic62 (Field1)

This dependence requires other deeper studies in this area. We note that if $RPTh80 \in (53, 86)$ then the text quartet estimation is the same. This indicates large intervals to have the same topic prediction of the text.

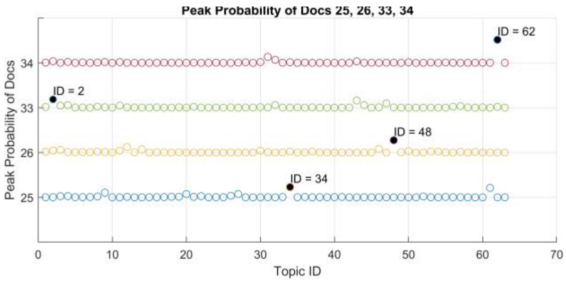


Fig. 16. Topic probability of the text quartet, Rigid LDA (docs 25, 26, 33, 34). The index of the peak is the ID of the topic.

For threshold $RPTh100$ only the largest probability counts in the decision. We name this classical case Rigid LDA (R-LDA). When the threshold is less than 100% we call it Flexible LDA (F-LDA) and the impact of it is explained in subsection E of this paper (see Fig. 16).

D. Impact of the text preprocessing on the LDA

We applied the algorithm to both the original and cleaned versions of the texts to assess the performance of the LDA method. Successfully, all 392 documents lacking fields were categorized into topics, with one or more fields being attached to each text.

Fig. 17 illustrates the topic assignment for the 392 documents in question. Each circle in the figure corresponds precisely to one topic assigned to the respective document. The chosen topic represents a list of individual fields that aligns with the feature vector patterns shown in Fig. 12 on the left. We observed a similar pattern in the allocation of predicted topics to documents categorized under pseudo-topics in both cleaned and uncleaned cases (refer to Fig. 17, left and right).

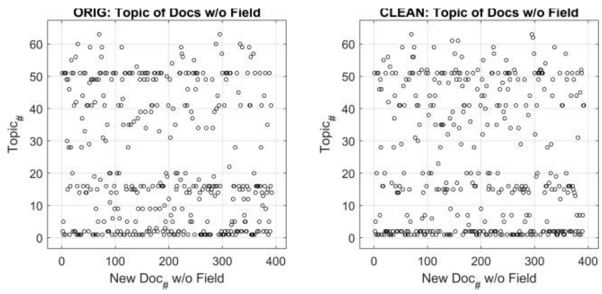


Fig. 17. Topics of Docs w/o Field (left); Topics of Clean Docs w/o Field (right).

To quantify dissimilarity, the Hamming distance between topics predicted for original and cleaned documents was assessed, revealing a range of [0, 7] (see Fig. 18). Large Hamming distances (greater than 5) occur very rarely. The majority of the distances are zero with the following mean, standard deviation and skewness values: $(\mu, \sigma, \gamma) = (1.059, 1.597, 1.453)$.

Our findings indicate that LDA exhibits a reduced sensitivity to the execution or omission of the cleaning task as a pre-processing step. To validate the effectiveness of this artificial cognitive capability-based method, we conducted manual testing by randomly selecting a few unlabelled legal texts. We then compared the automatic labelling performed by the LDA method to the actual categorization decided by a human evaluator.

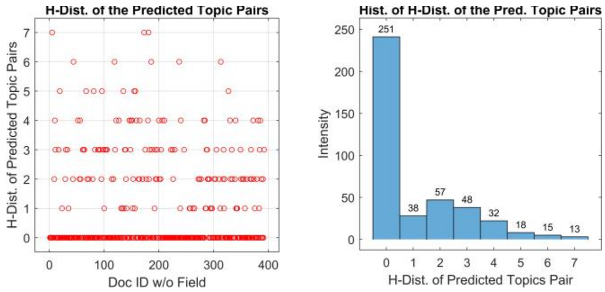


Fig. 18. Hamming distance of the topic prediction by Rigid LDA for cleaned and original documents (left); Histogram of Hamming distance of pairs of topics (right).

Despite a small sample size of fewer than ten entities which have a large deviation in their text length, we observed a correct matching ratio. The efficiency of the method was further demonstrated by the time it took for manual label assignment by the human evaluator. The manual assignment of one label consumed approximately one hour, considering the inherently time-consuming nature of reading and interpreting the text by a human.

E. Impact of the threshold value on the LDA

We extended the classical LDA by considering not only the largest probability value of the topics, but even others having a relative value close to it determined by the relative probability threshold $RPThx$ (%), $x \in [1, 100]$.

We assume that making decisions based on the topic group of top probabilities offers a more accurate prediction of individual labels compared to relying solely on the selected unique topic given by the top 1 topic probability. This pattern of individual labels constitutes a multilabel pattern, but it may differ from the list of topics determined by the teaching data set. This property has a greater impact when the top 1 topic has a low probability and is close to other topics.

We determined the number of significant topics for each unlabelled document. Mean, standard deviation and coefficient of variation metrics of the number of the significant topics vs. threshold $RPThx$ in the case of 392 unlabelled documents are presented in Fig. 19 (left) and in Fig. 18 (right), respectively.

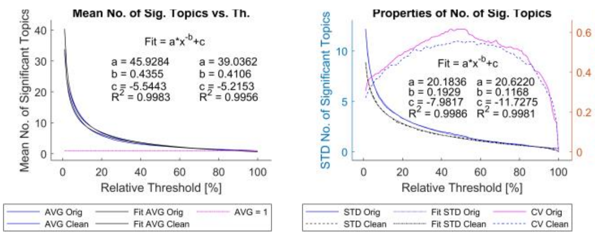


Fig. 19. Dependence of the number of significant topics vs. relative probability threshold (cleaned documents case): Mean (left); Standard deviation and coefficient of variation (right). In both figures, the first and second columns of the fit parameters (a, b, c) belong to the original and cleaned documents, respectively.

For both the original and cleaned unlabelled documents these metrics have a power-law dependence on the threshold value, demonstrating a remarkably high coefficient of determination ($R^2 > 99.5\%$). Fit parameter triplets (a, b, c) are provided in Fig. 18. Dependence of the coefficient of variation (CV) on the threshold $RPThx$ for both original and cleaned unlabelled documents is depicted in Fig. 18 on the right. We can observe that these curves are approximatively symmetric to the vertical axis at $RPTh50$, with a maximum value of 0.6. The impact of the text cleaning can be seen on the smoothness of the CV curves (refer to Fig. 18 on the right). The more the corpus is cleaned, the smoother the CV curve becomes.

IV. Conclusions

In the paper we exhibit automatic labelling methods based on multilabel analysis. The Latent Dirichlet Allocation method can be successfully applied to classify legal texts with multiple fields. Quantitative properties of the documents are influenced by the cleaning and normalizing of the texts. These pre-processing tasks will result in non-conformity to Zipf's law, because the intensity of words instead of inverse proportionality (exponent value = -1) of the rank order becomes a power function with exponent value = -0.63 in the case of the legal corpus. The LDA is less sensitive to the pre-processing of the texts by cleaning. The method in some situations offers the classification of the text to more than one similarly dominant topic. These similar topic pairs have low Hamming distance, causing the ambiguity of making labelling decisions. We implemented the Flexible LDA with enhanced properties, based on a single parameter to increase the efficacy of the labelling. More research is needed to analyse the relationship between the sensitivity of the LDA method and the state of text cleaning.

Acknowledgment

This work has been supported by the QoS-HPC-IoT Laboratory and project TKP2021-NKTA of the University of Debrecen, Hungary. Project no. TKP2021-NKTA-34 has been implemented with the support provided from the National Research, Development and Innovation Fund of Hungary, financed under the TKP2021-NKTA funding scheme.

References

[1] T. Váradi et al., "The MARCELL legislative corpus," in Proc. of the 12th Conference on Language Resources and Evaluation. LREC 2020. Marseille, France: European Language Resources Association, 2020, pp. 3761–3768.

[2] *Encyclopedia entry: laws in quantitative linguistics*. [Online]. Available: http://lql.uni-trier.de/ (2023-05-06)

[3] van C. J. Rijsbergen, Information Retrieval. 2nd ed. London: Butterworths, 1979. [Online]. Available: (2023-05-03) http://www.dcs.gla.ac.uk/Keith/Preface.html

[4] *Marcell Project Homegape*. [Online]. Available: (2023-02-22) https://marcell-project.eu/

[5] *EU Council Presidency Translator Workshop*. [Online]. Available: https://hr.presidencymt.eu//workshop (2023-02-23)

[6] M. Mohri, A. Rostamizadeh, and A. Talwalker, Foundations of machine learning. 2nd ed. Cambridge, MA: The MIT Press, 2018.

[7] J. Nam et al., "Large-scale multi-label text classification — revisiting neural networks," in T. Calders, F. Esposito, E. Hüllermeier, R. Meo Eds. Machine Learning and Knowledge Discovery in Databases ECML PKDD 2014, LNCS, vol. 8725, Heidelberg: Springer, 2014, pp. 437–452. DOI: 10.1007/978-3-662-44851-9-28

[8] J. Liu et al., "Deep learning for extreme multi-label text classification," in SIGIR'17: Proc. of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York, USA: ACM, 2017, pp. 115–124. DOI: 10.1145/3077136.3080834

[9] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet Allocation," *Journal of Machine Learning Research* vol. 3, no. 5, pp. 993–1022, 2003.

[10] M. D. Hoffmann, D. M. Blei, and F. Bach, "Online learning for Latent Dirichlet Allocation," in J. Lafferty, C. Williams, J. Shawe-Taylor, R. Zemel, A. Culotta Eds. Advances in Neural Information Processing Systems 2010, NIPS, vol. 23, Curran Associates, Inc., 2010, pp. 1–9.

[11] K. R. Canini, L. Shi, and T. L. Griffiths, "Online inference of topics with Latent Dirichlet Allocation," in Proc. of the International Conference on Artificial Intelligence and Statistics 2009, JMLR, vol. 5, Clearwater Beach, Florida: MLR Press, 2009, pp. 65–72.

[12] K. Johnson, Quantitative methods in linguistics. Malden, MA: Blackwell Publishing, 2008.

[13] S. Th. Gries, Quantitative corpus linguistics with R: a practical introduction. 2nd ed. New York: Routledge, 2016.

[14] *Homepage of the LF Aligner translator*. [Online]. Available: https://sourceforge.net/projects/aligner/ (2023-02-22)

[15] D. Varga et al., "Parallel corpora for medium density languages," in G. Angelova, K. Bontcheva, R. Mitkov, N. Nicolov, N. Nikolov Eds. Proc. of the RANLP 2005 (Recent Advances in Natural Language Processing), Borovets, Bulgaria, 2005, pp. 590–596.

[16] S. Duan, S. Chang, and J. C. Príncipe, "Labels, information, and computation: efficient learning using sufficient labels," *Journal of Machine Learning Research* vol. 24, pp. 1–35, 2023.

[17] G. Tsoumakas, and I. Katakis, "Multi-label classification: an overview," *International Journal of Data Warehousing and Mining* vol. 3, no. 3, pp. 1–13, 2007.

[18] J. Read et al., "Classifier chains for multi-label classification," *Machine Learning* vol. 85, pp. 333–359, 2011.

[19] M.-L. Zhang and L. Wu, "Lift: Multi-Label Learning with Label- Specific Features," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 1, pp. 107–120, 1 Jan. 2015, DOI: 10.1109/TPAMI.2014.2339815.

[20] M.-L. Zhang and Z.-H. Zhou, "A Review on Multi-Label Learning Algorithms, " in I*EEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 8, pp. 1819–1837, Aug. 2014, DOI: 10.1109/TKDE.2013.39.

[21] S. Hartman, "Derivational morphology in flux: a case study of word-formation change in German," *Cognitive Linguistics*, vol. 29, no. 1, pp. 77–119. 2018 DOI: 10.1515/cog-2016-0146

[22] M. Baroni and Stefan Evert, "Testing the extrapolation quality of word frequency models," *Proc. of Corpus Linguistics 2005*. [Online]. Available: https://www.birmingham.ac.uk/Documents/college-artslaw/corpus/conference-archives/2005-journal/Lexiconodf/EvertBaroni2005.pdf (2023-07-12)

[23] P. Baranyi and Á. Csapó, "Definiton and synergies of cognitive infocommunications," *Acta Polytechnica Hungarica*, vol. 9, no. 1, pp. 67–83, 2012.

[24] P. Baranyi, Á. Csapó, and G. Sallai, Cognitive infocommunications (CoginfoCom). Springer, 2015.

[25] I. Horváth et al., "Definiton, background and research perspectives behind 'Cognitive Aspects of Virtual Reality' (cVR)," *Infocommunications Journal*, Special issue: Internet of Digital & Cognitive realities, pp. 9–14, 2023. DOI: 10.36244/ICJ.2023.SI-IODCR.2.

Erzsébet Tóth is an assistant professor at the Faculty of Informatics, University of Debrecen, Hungary. She had a degree in English language and literature and library and information science in 1995. In 2008 she obtained her Ph.D degree in information science and technology. In the doctoral dissertation she investigated the evaluation of search engine performance. She is involved in a virtual library project that focuses on the presentation of digitized library content in three-dimensional space, and she studies the enhanced possibilities of the English language teaching and learning in virtual learning environment. Since 2022 she has been an IEEE member.

Zoltan Gal is currently a full professor at the Faculty of Informatics, University of Debrecen, Hungary. He earned MSc in electrical engineering and computer science from the Technical University of Timisoara, Romania and PhD in informatics sciences from the University of Debrecen. His scientific interest is focused on distributed processing and communication systems, sensor technologies and services in the Internet of Things.

He was the CIO of his institute for 20 years and developed the university-level metropolitan area highs-speed data network and services with over 10k Internet nodes. He is Cisco Certified Network Professional lecturer since 1999 and taught over five hundred network professionals in the field. Starting in 2015 he is head of the Centre of High-Performance Computing at his university. He is an IEEE member since 1996 and published over one hundred twenty scientific conferences and journal papers: He supervises his own R&D&I project called QoS-wwHPC-IoT Laboratory.

Knowledge Base Development for Second Language Learning in the 3D Virtual Space

István Károly Boda, and Erzsébet Tóth, *Member, IEEE*

***Abstract*—In our study first we provide a short overview of the 3D virtual library project which started about ten years ago as part of the Cognitive Infocommunications (CogInfoCom) research. The current implementation of the virtual library model exploits the 3D features of the MaxWhere Seminar System. In our study we would like to summarize the classroom experiences that we brought together in teaching English as a second language for students of computer science majors at the Faculty of Informatics, University of Debrecen in the academic year of 2020 and 2021. Our main purpose was to improve the students' linguistic knowledge and to collect their opinions and suggestions about the content of the learning material of our virtual library and where they think is necessary to modify it. Summarizing their views, we decided to add further vocabulary items, contexts and explanations to the material. In addition, their achievements proved that successful language learning needs carefully prepared tests and exercises which support self-assessment and increase motivation. In general, the more tests are available the more efficient the learning process is. But preparing good and varied tests manually is a relatively slow and exhausting work. Therefore, we intended to use JavaScript technology to develop an algorithm which can generate tests and exercises automatically, based on the knowledge base of the virtual library.**

***Index Terms*—automatic test and exercise generation, CogInfoCom, MaxWhere Seminar System, second language learning, three-dimensional (3D) virtual library project, virtual learning environment**

I. Introduction

ABOUT ten years ago we launched a *virtual library project* as part of the cognitive infocommunications (CogInfoCom) research [1-2], focusing on the presentation of selected library content in virtual 3D environment the features of which have been thoroughly investigated by several CogInfoCom studies since then (e.g. presentation of virtual rooms and buildings in the 3D space [3-4], developing and using effective 3D learning environment [5-6], cognitive and psychological aspects of the 3D environment [7-8] etc.). The main objective of the virtual library project was to collect, organize and present relevant verbal and multimedia content in the 3D space about the ancient Library of Alexandria and classical Greek literature translated into English (e.g. texts about the life and work of Callimachus, translations of selected literary texts from the works of prominent ancient authors etc.) [9-11]. Although the 3D virtual library is designed for various purposes, *language learning* has proved to be the most promising application of the collected virtual library content [12-13] because of its potential usefulness and the importance of advanced language skills (in our case, in English) to convey the ancient cultural heritage reflected in the collected material to today's cultural environment.

The literary and cultural heritage of antiquity and especially the collection of the ancient Library of Alexandria are usually referred to as a symbol of universal human knowledge and wisdom. Its catalogue covered all Greek literary works which were once available in its holdings, and the catalogue was used as a bibliography of ancient Greek literature for centuries. The 3D virtual library model of our project (3DVLM) is based on the classification system of the ancient Library of Alexandria, especially on the famous Pinakes invented by the great scholar-poet Callimachus in the 3rd century BC [12].

The current implementation of the 3D virtual library model makes use of the innovative 3D environment of the MaxWhere Seminar System [14] the features of which, and especially the embedded web browsers called smartboards, fully support the implementation of the basic concepts of the 3DVLM [15-17]. The core content of the knowledge base of the virtual library focuses on the classical heritage the European culture is based on. Our main idea is to revive and convey the message of ancient times to the present-day culture, which seems to be really crucial with respect to the young social generations. We firmly think that this mission can be best accomplished by language learning which, along with its obvious importance in our internet-based society, can serve as a bridge between the ancient times and the "modern" culture of the XXI. century. The basic idea of the virtual library project is that with a carefully elaborated way and methodology the ancient thoughts and values can effectively be translated (both literally and figuratively) to the young members of the generation CE (i.e. the generation cognitive entities whose members are growing up in "entangled co-evolution with ICT" [2])

The index page of the current implementation of the virtual library project begins with a 2D map of the virtual library which shows and briefly explains the main entry points to the content of the library (Fig. 1).

István Károly Boda is with the Department of Mathematics and Informatics, Debrecen Reformed Theological University, Debrecen, Hungary (e-mail: boda.istvan@drhe.hu).
Erzsébet Tóth is with the Faculty of Informatics, University of Debrecen, Debrecen, Hungary (e-mail: toth.erzsebet@inf.unideb.hu).

DOI: 10.36244/ICJ.2024.5.8

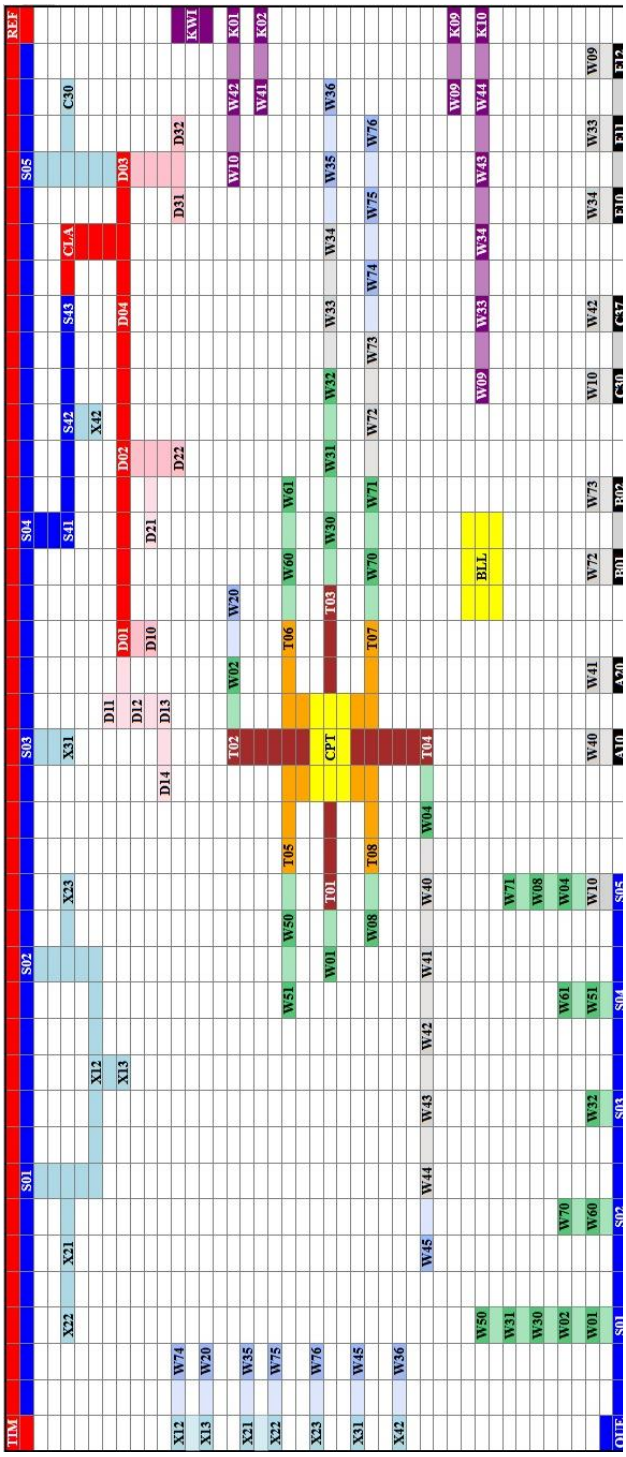


Fig. 1. The 2D map of the current implementation of the virtual library [xix]. Each abbreviation in the map refers (and is linked) to a corresponding part of the virtual library; e.g. clicking on [T06] (about in the centre) we can go to the thesaurus page described in detail in Section III.

In the following section a brief account will be provided how we were teaching English as a second language for the students of the Faculty of Informatics, University of Debrecen on the basis of the preprocessed texts and supporting material delivered by the knowledge base of our virtual library.

II. A SUMMARY OF CLASSROOM EXPERIENCES

In the academic year of 2020 and 2021 we could teach the compiled learning material of our virtual library for students of the Faculty of Informatics, University of Debrecen at Bachelor level. Our main objective was to develop the students' linguistic competence and to collect their opinions and suggestions about the content of the learning material (e.g. where they think is necessary to correct or modify it, where to add further comments, vocabulary items, explanations etc. to it). In the autumn semester an optional university course was offered for the students in a blended learning form. The only admission criterion for students to this course was to have a successful placement test which measured their skills of using grammatical rules properly in different contexts at B2 level. Using this test, we could quickly select those students from the 120 candidates for the course whose language skills were satisfactory (i.e. at upper-intermediate or advanced level). Finally, we had 60 students altogether who were divided into three different groups for teaching. Although in the beginning we could deliver our courses in a traditional way where a face to face contact was possible between teachers and students at this time of the semester, later, because of the world-wide pandemic, we had to change the way of teaching from traditional to online starting from 9th November, 2020.

Below we provide a short overview of the activities we planned, the good experiences and practices we had, and the main difficulties we had to deal with during the courses. We **emphasize in bold the most common problems** we faced and the suggestions or observations the students made to handle them.

A. Teaching environment and the use of information technology

The lessons for the two groups of students were placed in the same big lecture hall where we could use only a projector and the blackboards as teaching aids. But the lessons for the third group were located in a smaller classroom which was equipped with the same teaching aids. In this traditional learning environment, we decided to present the two-dimensional version of the learning material for the students which was the same as the three-dimensional one available in the 3D Castle space of the MaxWhere VR platform.

B. Texts taught during the course

According to the time schedule of the course syllabus, during 6 weeks we could focus on the learning material about Callimachus and the Library of Alexandria in two teaching hours (i.e. 2*50 min) per week. Below we provide the list of the primary and secondary texts from our virtual library that we were teaching for the students during the lessons. (The codes that identify the texts can be seen in Fig. 1).

- Callimachus [S01],

  Cyrene, Libya [X12],

  The Dorians [X13],

  The Ptolemaic dynasty [X21],

  The Ptolemaic Kingdom [X22],

- Ancient Greek Literature [CLA],
  The nine lyric poets [D21],
- Callimachus in Alexandria [S02],
  Alexandria [X23],
- The Great Library of Alexandria [S03],
  The Mouseion [X31],
- The Pinakes [S04],
  The collection of the Great Library [S41],
  Callimachus' classification system [S42],
  Binomial nomenclature [X42],
  The structure of the Pinakes [S43],
- The works of Callimachus [S05].

C. Text comprehension task

In the case of the S01, X12, X13, X21, X22 content units we could practice text comprehension by answering to various questions which were closely related to the content of the given texts. In the text comprehension tasks we perceived that the students could easily follow and understand the main topic of the complex texts in general. However, they often failed **to explain the precise meaning of unknown words or phrases, or to recognize new verb or collocation patterns** emerging in various contexts.

D. Vocabulary Teaching

We often made use of the vocabulary items integrated into the learning material. Those items are presented for the students in order to help them to understand better and memorize the **spelling, pronunciation and definition of selected words** and their related term(s) as well. While we were reading and processing the texts about Callimachus, the students were asked **to mention or provide a synonym or antonym of, to define, or paraphrase the meaning of selected words or expressions** [18]. We learned that it was the most difficult task for them. In case we explored a new word, expression or collocation in the text we were studying with the students, we suggested that they should learn and memorize it 'as is'.

E. Collecting opinions from the students

We distributed a small paper among students towards the end of the lessons to collect their opinions or proposals about the content of the learning material by posing them the following question: *"Which words, collocations or grammatical structures caused you any difficulties in the learning material?"* According to these feedbacks the students had a big **problem with the pronunciation and spelling of proper names** occurring in the texts. For example, they asked us to add new vocabulary entries to the learning material about the spelling and pronunciation of the name of the major ethnic groups in Classical Greece (not to mention historical figures, mythological characters, geographical names etc.).

F. Preparation for the lessons

In advance, we prepared a lot for these English lessons and looked for the precise *definition* of each new word and its most frequently used collocations in a monolingual dictionary [e.g. 19] which seemed to be relevant in the context. In several cases we also chose *sentence samples* or, in general, *concordances* (from various sources, e.g. from literary works, monolingual or quotation dictionaries etc.) which conveyed the various shades of meaning of the given word in different contexts. There were cases when we had to provide further explanations for the students about the meaning of *special terms*, e.g. those used for describing historical periods, art periods etc. ("Archaic Greece", "Old Comedy", "New Comedy", "Hellenistic Period" etc.) From time to time, we also checked the *grammar* [22] used in the studied texts if further grammatical explanations are required for the students.

G. Integration of concordances into teaching

Concordances can provide a simple and valuable help to teach real English in the classroom [20-21]. We selected a lot of concordances that we thought might be helpful for the students, and organized them around specific keywords in separate *thesaurus pages*. Although it is not very easy to find the best method for integrating concordances successfully into teaching, we presented the content of the compiled thesaurus pages to the students and explained the page structure for them putting special emphasis on the importance of concordances. In case a concordance had an obscure or ambiguous part (with respect to meaning, grammar, style, imagery etc.) we provided special *explanatory notes* for the given concordance in the same format which we used for the main content units of the learning material. Note that the notes often contain references to other concordances and explicit concordances as well.

The explanatory notes attached to certain phrases in the S05 content unit proved to be very useful for the students. Using these notes, the students could understand better those collocations which included abstract concepts such as „*an aesthetic of smallness and perfection*", „*his expression of what constituted excellence*", „*drive their wagons*" etc. In these explanatory notes we used concordances in order to support the language learners to understand and memorize better the meaning of certain collocations and broaden their vocabulary at the same time. Note that only at this point of the learning material we could easily integrate concordances into teaching. As a consequence, we found that concordances, which play an important role in activating language, would principally be associated with a given context.

H. Assessment of students' language competence

After 6 teaching weeks the students wrote a test paper on the vocabulary and grammar of the learning material in a traditional form. The so-called "word test" contained new words, collocations, parts of sentences or whole sentences that the students had to translate from English to Hungarian and vice versa. They had to provide English definitions of five terms that were selected from the studied texts of the learning material. The students found this exercise a bit difficult, but they could cope with it successfully. In addition, they had to solve a grammatical exercise related to advanced English grammar [22]

The students found this exercise a bit difficult, but they could cope with it successfully. In addition, they had to solve a grammatical exercise related to advanced English grammar [22] which occurred in the studied text about ancient Greek literature [CLA].

On the whole, we can conclude that the majority of the students could acquire successfully the vocabulary and the grammar of the primary and secondary texts of the virtual library because a relatively small proportion (15%) of them received unsatisfactory grade for their test paper. (Note that those students who failed at first could later try again to improve their results.). Due to the relatively small number of students we consider the students as one homogenous group and did not examine the test results according to different aspects. As regards the further assessment of the learning material, we plan to use the Google Translate service [26] as an AI agent that can be metaphorically considered as a virtual language learner at an intermediate (or advanced) level. GT can translate selected parts of the learning material and the evaluation of the results can provide unlimited number of empirical data which we can use continuously to assess and improve our learning material.

On the basis of the students' opinions and suggestions (e.g. on the spelling and pronunciation of proper names, the explanation of special terms etc.) we intended to **add further vocabulary items and explanations to our learning material** to make it more usable and understandable for those language learners who use our material to improve their English (preferring either classroom or online learning environment, or independent learning).

III. Content development for classroom use

On the basis of our classroom experiences we decided to further develop the knowledge base of the virtual library in order that it could be much more effectively used in language learning and teaching. Because the use of thesaurus pages, and especially concordances caused considerable difficulties for students we focused on the full revision and development of the structure and content of **thesaurus pages**. In the following, we would like to describe a selected page in detail on the one hand, and explain the applied page elements and tools on the other hand.

Thesaurus pages are organized around given *microcontexts*, where the term 'microcontext' means a group of semantically related words (i.e. synonyms, antonyms, related or contrasted words) that fits into a given grammatical or collocation pattern. We present the words in a way following the structure of well-known thesauri, synonyms dictionaries and lexicons [i-vi] where most of the related words were selected from (creating several subgroups, sometimes in two or three hierarchical levels, according to the meaning of the words). The collocation patterns, in turn, can be found in collocations dictionaries [vii]. The microcontexts are illustrated by carefully selected concordances, similarly to the example sentences of the entries of monolingual dictionaries [e.g. vii-xii] where most of the concordances were selected from, although we used production and quotation dictionaries as well [xiii-xv] and various literary and other texts (e.g. [xvi-xvii]). Here we use the general term 'concordance' because in our case concordances include not only selected sentences, but quotations or short extracts from various texts as well which we think can be useful for language learners to develop their vocabulary and related skills (e.g. grammar, composition, style etc.).

We selected the thesaurus page entitled [T06] which is related to the keywords *account, description, report; narrative, story; article, document, essay, review, study* etc. The page is organized around the collocation pattern [**adj**+noun] where the possible nouns (i.e. the keywords which can be placed into the "noun" part of the pattern) can be seen in Fig. 2.

[G154] paper, *document*, *record*, register; minutes, notes;
[G155] letter, note, label;
[G156] notebook, excercise book, folder, file;
[G157] *account*, *description*, *report*, version; *testimony*;
[G157] *story*, *tale*, *anecdote*; *narrative*;
[G158] *essay*, composition, *paper*; (piece of) *work*, *study*, *survey*;
→ [G65] speech, talk; lecture; presentation;
[G159] *article*, feature, sketch, editorial, headline; *review*;
[G160] material, copy, print;
→ [G242] *text*;
[G161] summary, synopsis, abstract, abridgement, précis, brief;
[G162] book, manuscript, scroll, script, draft, textbook, manual, diary;

Fig. 2. The list of keywords used in thesaurus page [T06].

Some of the possible adjectives (which can be placed into the "adj" part of the pattern) can be seen in Fig. 3. For example, valid combinations are *"accurate account"*, *"apt description"*, *"comprehensive report"* etc.

*[**adj** + noun]; [noun + BE + **adj**]*
- *[CONTENT]*
  - *accurate, confirmed, correct, exact, legitimate, precise, proper, right*
    - *careful, conscientious, expert, meticulous, painstaking, professional, scrupulous, thorough*
    - *appropriate, apt, fitting, relevant, suitable*
    - *clear, coherent, comprehensible, intelligible, plain, straightforward, unambiguous*
    - *abstract, complete, comprehensive, complex, detailed, erudite, extensive, full, general, in-depth, systematic*
      - *entire, whole*
    - *consummate, excellent, fine, perfect, superb, supreme*

Fig. 3. The first part of the list of adjectives used in thesaurus page [T06].

At the end of the thesaurus page [T06] all the available combinations are listed in two tables. In the first table the collocations are ordered by nouns (e.g. 'account', 'anecdote' etc.; see Fig. 4), and in the second table the collocations are ordered by adjectives (e.g. 'accurate', 'ancient' etc.). Note that currently there are 18 nouns (including plural forms) and 221 adjectives which make possible, at least theoretically, 3978 combinations of the corresponding words according to the given collocation pattern. In the thesaurus page there are currently 58 combinations that are validated (and illustrated) by selected concordances.

COLLOCATION PATTERNS (ORDERED BY NOUNS)

adj	noun	frequency
account, accounts		
accurate	account	1
contradictory	accounts	1
detailed	account	2
detailed	accounts	1
different	accounts	1
firsthand	account	1
vivid	account	1

Fig. 4. The collocations of the word 'account' listed in the first table at the end of the thesaurus page [T06].

In the thesaurus page we arranged and grouped the selected concordances according to the adjectives used in the applied collocation pattern (formatted in bold font type). In the concordances the keywords were emphasized using small-caps font variants. In some cases we presented the context of the concordance as well, inserting a short text immediately after the concordance. Where a rare or possibly "difficult" word occurred in the concordance (denoted by an asterisk) we added a link to a vocabulary entry at the end of the thesaurus page explaining the meaning of the given word (see Fig. 5.1-5.2).

[accurate]
accurate description = a description which is correct, exact and precise in all details

- *At a committee meeting we need to have an **accurate** ACCOUNT recorded* of our discussions.*
  *Have you ever been to a **committee meeting**? An important thing has to be settled at the very outset: who is going to take the minutes? We need to have an accurate account recorded of our discussions and sadly our memories are not always that reliable so we have someone take minutes. These can later be consulted so that we can follow through on the decisions we've taken and appropriate action can be taken.*
- *Quantum mechanics has to be used to provide an **accurate** DESCRIPTION of the microscopic system.*
- *A more **thorough**, **accurate** DESCRIPTION of the manuscript is under preparation.*
- *He has always kept an **accurate** RECORD of his spending.*

Fig. 5.1. The concordances containing the adjective *'accurate'* in the given collocation pattern.

[4] *record* [rɪ'kɔ:d] an account or facts = to set down* an account (of a meeting etc.) or facts (about historical events etc.) in writing or some other permanent form for later reference

- *An official report **records** that at least half the nation's monuments are in need of repair.*
- *At a committee meeting we need to have an accurate account **recorded** of our discussions.*
- *An authentic document **records** that the battle took place six years earlier.*

Fig. 5.2. The vocabulary entry explaining the verb *'record'* by a short definition and several concordance samples.

In the learning process it is strictly necessary for a student to understand the structure and aim of the learning material in order to (repeatedly) read and memorize its content. But to memorize almost 60 concordances certainly needs additional support. Therefore we created a semantic map of the most important collocation patterns and concordances which shows the semantic relationships of the patterns and helps visually memorize them (Fig. 6).

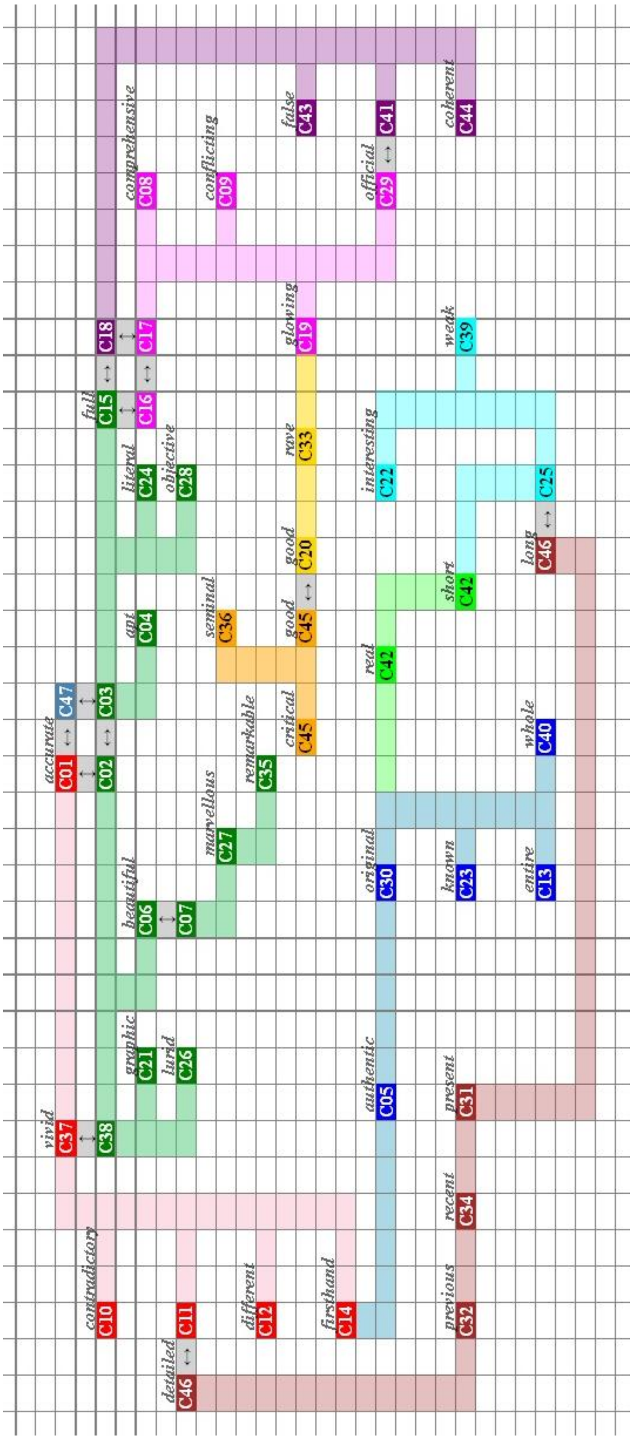


Fig. 6. The semantic map of the most important collocations represented in the thesaurus page [T06]. Each abbreviation refers (and is linked) to a specific concordance in the page; e.g. clicking on [C46] (labeled with the adjective 'long') we can go to the concordance "*He made a **long** and detailed study of how animals adapt to their environment.*" [x].

Note that we deliberately used different colors for different keywords in the patterns (e.g. red for 'account', green for 'description', etc.). We linked the patterns which contained either the same or syntactically related adjectives (i.e.

synonymous / antonymous or related / contrasted words as in thesauri), e.g. ‘beautiful’ ~ ‘marvellous’ ~ ‘remarkable’; ‘authentic’ ~ ‘original’ ~ ‘real’; ‘long’ ~ ‘short’ etc.

Finally, we created 50 **tests** for the students to practise their skills and improve the knowledge they have acquired. Note that doing tests repeatedly is also a very effective way of *learning and memorizing* the content provided by the tests.

Every test is based on a given concordance but the adjective and noun parts of the concordance are replaced by 4-4 randomly generated words in addition to the “valid” words which provide the right solution of the test (thus a student should choose the correct words from 5 options for each variable part of the sentence). So there are fixed and variable parts of a given sentence in a test, and sometimes, for either grammatical or didactic purposes, there are more than two variable parts (e.g. a student should choose the right article, the correct form of the verb etc.). We tried to create the algorithm which generates the random tests that as many combinations appear to be more or less acceptable as possible. We can always find the offered solution of every test clicking on the OK button at the end of the sentence (see Fig. 7-8).

1. TEST

At a committee meeting	*we need to have*	*– a an*	*accurate comprehensive fair inadequate introductory*	*account anecdotes essay story works*	*recorded of our discussions.*	***ok***

Fig. 7. The test based on the first concordance which can be seen in Fig. 4.

1. TEST

At a committee meeting	*we need to have*	*– a* ***an***	***accurate*** *comprehensive fair inadequate introductory*	***account*** *anecdotes essay story works*	*recorded of our discussions.*	***ok***

Fig. 8. The solved test which can be seen in Fig. 7.

Note that, in addition to grammar rules, we need to get realistic alternatives for the words which appear in the variable parts of the sentences. In this respect, the semantically related nouns (keywords) and adjectives are crucial. But we cannot be quite sure that all the combinations are correct unless we have a real occurrence which is provided by one or more concordances (which validate a possible combination of words).

IV. Presentation of the Virtual Library Content in the 3D Space

As mentioned before, the current implementation of the 3DVLM exploits the excellent 3D features of the MaxWhere Seminar System including, in the first place, the effectively and decoratively arranged *smartboards* in ready-made 3D virtual spaces where the main content (e.g. *cabinets,* corridors as well as cabinet walls) and the navigation / organization tools (index, thesaurus, reference etc. *pages*) of the virtual library [16–17] can be presented. There are a lot of well-designed and spectacular 3D virtual spaces available on the MaxWhere site [14] which can be used in almost every context, although each space has its unique and distinguished features. For the arrangement and presentation of the virtual library content, in our previous publications [16–17,21] we chose the 3D Castle virtual space. Due to the flexibility of the 3DVLM, we can use other 3D spaces as well. For the new implementation we decided to try and use another 3D space, namely the 3D Library virtual space which contains its smartboards in a two-storey virtual library building. In the following, we would like to present some screenshots, along with detailed explanations, which would illustrate how to access selected virtual library content in the 3D Library space.

Let the starting point be the navigation page [17] (Fig. 9).

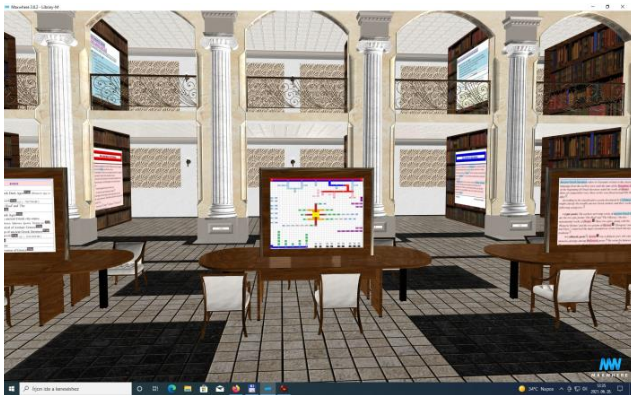


Fig. 9. A screenshot focusing on the *navigation page* of the virtual library located on the ground floor in the MaxWhere 3D Library space.

In the foreground of the image presented in Fig. 9 we can see smartboards which jointly serve as an “information desk” of the 3D library. They offer “smart” access to the navigation / organization tools of the virtual library:

– the *navigation page* is located at the centre of the image; it contains the 2D map of the virtual library where the primary and secondary texts, and other content units are identified by corresponding codes (cf. section II/B. and Fig. 1);

– on the left side we can see a small part of the page which provides a *timeline* of some historical milestones of the ancient era;

–on the right side a part of the *category page* [17] can be seen which contains descriptions of the main classification categories and shows their hierarchical arrangement.

In the background of the image presented in Fig. 9 we can observe some additional smartboards as follows:

– the smartboards on the ground floor of the 3D library present the content of the *main cabinets* of the virtual library containing primary texts about Callimachus (see later);

– the smartboards on the first floor of the 3D library present the *thesaurus pages* of the virtual library including the thesaurus page [T06] introduced in detail in Section III.

Because the main function of the information desk is to help the users find the relevant information, we placed the content of the navigation / organization pages on the wall of the 3D library as well (Fig. 10). Here at the centre of the image we can see again the 2D map of the virtual library which provides entry points to the texts and other content units of the library. Note that as the collection of the virtual library grows new codes and links will appear in the map.

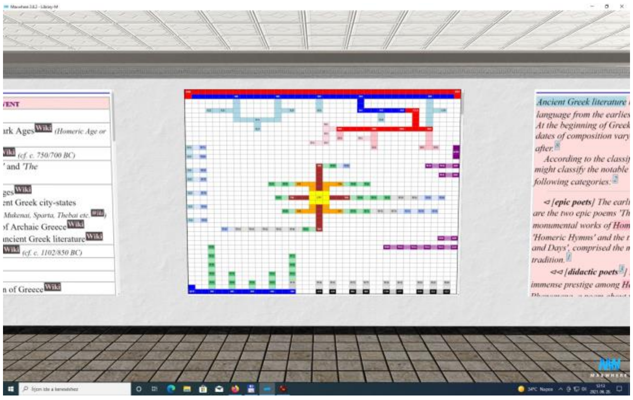


Fig. 10. A screenshot of three *navigation / organization pages* of the virtual library located on the wall in the MaxWhere 3D Library space.

The content of the main cabinets is organized around selected slides about the life and work of Callimachus (labeled as Callimachus / S01, Callimachus in Alexandria / S02, The Great Library of Alexandria / S03, The Pinakes / S04, The works of Callimachus / S05 etc. [16–17,23]). They can be found on the ground floor of the 3D library just behind the information desk. The slide about Callimachus, and that about the Great Library of Alexandria can be seen in Fig. 11.

Fig. 11. A screenshot of the *cabinets* which contain slides about Callimachus and the Library of Alexandria located on the ground floor in the MaxWhere 3D Library space.

From a different angle, the slide about the Pinakes can be seen in Fig. 12.

Note that the slides about the life and work of Callimachus can serve as the starting point for the language learners to the full content of the library. After them they can explore other texts some of which are directly or indirectly connected to the primary texts about Callimachus and his work. For example, the presented text about Pinakes, which "was one of the first known documents that lists, identifies, and categorizes a library's holdings" [xvi], involves both a selected text about the collection of the Great Library of Alexandria [S41] and a brief summary about Callimachus' classification system [S42].

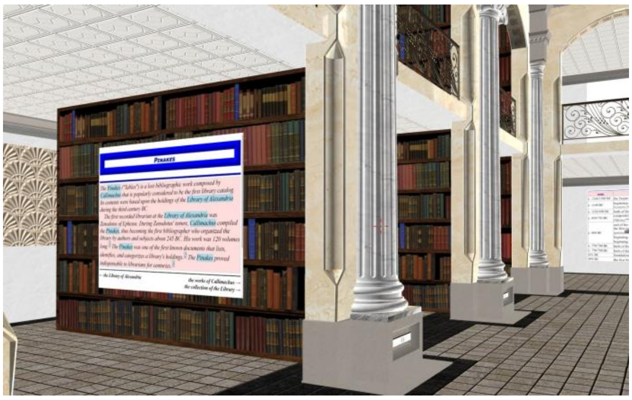


Fig. 12. A screenshot of the *cabinet* which contains a slide about the Pinakes located on the ground floor in the MaxWhere 3D Library space.

For those who are interested in the content of the virtual library that we have presented in this study (including the texts, the thesaurus pages, additional supporting materials etc.), the current implementation of the virtual library project can be accessed in 2D web page form via the internet [xviii-xix].

V. Conclusion

Obviously, the linguistic content which the thesaurus page [T06] represents covers only a small piece of the practically infinite variations the natural language can offer. Nevertheless, we think that it can teach and broaden very important skills for the students, e.g.

– the semantically related words (either the adjectives or the keywords), as well as the additional words which occur in the concordances, considerably improve the students' vocabulary;
– the students can get "ready-made" patterns which they can apply directly, and they learn the way how such patterns can be created;
– some concordances (e.g. quotations) and/or their attached context convey useful and valuable meanings which are well worth memorizing;
– hopefully the structure of the material motivates the students to observe and collect similar patterns.

In order to support language learners, we developed a knowledge base the content of which we try to gradually improve with additional items and tools that might be useful for language learners (e.g. preprocessed texts, dictionary and encyclopedia entries, maps, elaborated microcontexts, selected concordances, quotations, selected passages from literary works, generated tests etc.). Because the arrangement of the various items and their relationships play a significant role in the learning process, we took full advantage of both the hypertext-based 2D and the virtual 3D environment *mapping and visualizing* the compiled material of the virtual library using the MaxWhere Seminar System with its excellent features [23].

Because "cognition, metacognition, procedural skills, and motivational factors are important determinants of learning activity" [27], cognitive aspects of the virtual library model are of utmost importance in language learning. In this respect, we can differentiate several levels, such as

– language level which, among others, makes possible to build words, expressions (phrases, idioms etc.) and sentences using grammatical (e.g. collocation or verb) patterns, stylistic and rhetorical figures or devices etc. (also called syntactic level [28]);

– textological level which, among others, makes possible to build texts (also called semantic level [28]);

– intertextual or hypertextual level which, among others, makes possible to build coherent (or semi-coherent) texts from different texts (also called pragmatic level [28]);

– presentation level which, using interactive 2D browsers and/or various 3D VR or AR (Virtual or Augmented Reality) tools, presents grammatical, textual, intertextual, hypertextual etc. relationships between the lexical items or units of the presented knowledge base and makes possible to have access to them in the virtual 2D or 3D space.

Note that the higher levels are based on the lower ones, and the resulting complex structure is one of the key factors that contribute to the overall effectiveness of the learning process. With regard to the fourth level, cognitive infocommunications research shows that using digital 3D VR and AR technology in education can support cognitive processes such as finding, processing, memorizing, and recalling information. Moreover, 3D environments are also “capable of providing users with a much higher level of comprehension when it comes to sharing and interpreting digital workflows” [29], e.g. when using e-learning tools or participating in a collaborative learning process [30].

At the first level the virtual library offers various solutions which can support the language learning process. As we have seen above, a thesaurus page is organized around separate but interwoven elements of knowledge called microcontext (i.e. two separate groups of semantically related words, illustrated by a list of selected concordances). We added tests to help the learners memorizing vocabulary and lexical items which seem to be necessary for them “to become a long-term part of the learner’s own store of English” [24]. The distinction between learning lexical items and their context by reading the pages repeatedly and practicing them by using tests raises the question whether it is worth presenting those parts of the knowledge base simultaneously, in separate smartboards. For example, one smartboard can display the tests, and another (practically the one that is next or opposite to the other one) can display the vocabulary, collocations, concordances etc. which the tests are based on. It can be a real benefit which only the 3D environment can provide.

As feedbacks are always important, future works could focus on evaluating the syntactic and semantic level using AI tools (e.g. the Google Translate service) and/or assessing the motivation and performance of language learners when using the 2D and 3D forms of the virtual library.

Acknowledgment

The results presented in this paper have partially been achieved in the Virtual Reality Laboratory of the Faculty of Informatics of the University of Debrecen, Hungary. This work has been supported by QoS-HPC-IoT Laboratory and project TKP2021-NKTA of the University of Debrecen, Hungary. Project no. TKP2021-NKTA-34 has been implemented with the support provided from the National Research, Development and Innovation Fund of Hungary, financed under the TKP2021-NKTA funding scheme

Sources

[i] Webster's Collegiate Thesaurus. Springfield, MA: G. & C. Merriam Company, 1976.

[ii] Roget’s Thesaurus of English Words and Phrases. (rev. by Dutch, Robert A.,) Penguin Books, 1978.

[iii] Roget’s Thesaurus of English Words and Phrases. (ed. by Kirkpatrick, Betty) Penguin Books, 1987.

[iv] The Penguin Dictionary of English Synonyms and Antonyms. (ed. by Fergusson, Rosalind,) Penguin Books, 1986.

[v] Oxford Learner’s Thesaurus. A Dictionary of Synonyms. Oxford: Oxford University Press, 2012.

[vi] Tom McArthur, Longman Lexicon of Contemporary English. Longman, 1984.

[vii] Oxford Collocations Dictionary for Students of English. Oxford University Press, 2002.

[viii] Longman Dictionary of Contemporary English. Longman, 2001.

[ix] Longman Dictionary of American English. Pearson-Longman, 2009.

[x] Hornby, A. S., Oxford Advanced Learner's Dictionary of Current English. Oxford: Oxford University Press, 2004.

[xi] Macmillan English Dictionary for Advanced Learners. Oxford: Macmillan Education, 2002.

[xii] Cambridge International Dictionary of English. Cambridge University Press, 1995.

[xiii] Longman Language Activator. The World’s First Production Dictionary. Longman, 1993.

[xiv] The Oxford Dictionary of Quotations. Oxford: Oxford University Press, 1964.

[xv] Bartlett, John, Familiar quotations. A collection of passages, phrases, and proverbs traced to their sources in ancient and modern literature. Boston-Toronto-London: Little, Brown and Company, 1992.

[xvi] *Wikipedia*.
[Online]. Available: https://en.wikipedia.org/wiki/Callimachus etc. (2019-07-15)

[xvii] *Project Gutenberg*.
[Online]. Available: http://www.gutenberg.org/ (2020-06-04)

[xviii] *3DVLM | thesaurus [T6]*.
[Online]. Available:
https://bodaistvan.hu/callimachus/_thes/accurate_description.html (2023-04-29)

[xix] *Interactive map of the three dimensional virtual library (3DVLM).*
[Online]. Available: https://bodaistvan.hu/callimachus/map.html (2021-06-29, 3D version; 2023-04-28, 2D version)

References

[1] P. Baranyi and Á. Csapó, “Definition and synergies of Cognitive Infocommunications,” *Acta Polytechnica Hungarica*, vol. 9, no. 1, pp. 67–83, 2012.

[2] P. Baranyi, A. Csapó, and Gy. Sallai, Cognitive Infocommunications (CogInfoCom). Springer International, 2015.

[3] A. Gilányi, A. Rácz, A. M. Bólya, J. Décsei, K. Chmielewska, “A Presentation Room in the Virtual Building of the First National Theater of Hungary,” in P. Baranyi Ed. CogInfoCom 2020. Proc. of the 11th IEEE International Conference on Cognitive Infocommunications, (Mariehamn, Finland, 23-25 Sept 2020,) pp. 519–523.

[4] A. Rácz, A. Gilányi, A. M. Bólya, J. Décsei, K. Chmielewska, “On a Model of the First National Theater of Hungary in MaxWhere,” in P. Baranyi Ed. CogInfoCom 2020. Proc. of the 11th IEEE International Conference on Cognitive Infocommunications, (Mariehamn, Finland, 23- 25 Sept 2020,) pp. 575–576.

[5] I. Horváth, "How to Develop Excellent Educational Content for 3D VR," in P. Baranyi Ed. CogInfoCom 2019. Proc. of the 10th IEEE International Conference on Cognitive Infocommunications, (Naples, Italy, October 23-25, 2019,) pp. 483–489.
[6] I. Horváth, "Personalized Learning Opportunity in 3D VR," in P. Baranyi Ed. CogInfoCom 2020. Proc. of the 11th IEEE International Conference on Cognitive Infocommunications, (Mariehamn, Finland, 23-25 Sept 2020,) pp. 425–439.
[7] B. Berki, "Sense of Presence in MaxWhere Virtual Reality," in P. Baranyi Ed. CogInfoCom 2019. Proc. of the 10th IEEE International Conference on Cognitive Infocommunications, (Naples, Italy, October 23-25, 2019,) pp. 91–94.
[8] B. Berki, "Navigation Power of MaxWhere: a Unique Solution," in P. Baranyi Ed. CogInfoCom 2020. Proc. of the 11th IEEE International Conference on Cognitive Infocommunications, (Mariehamn, Finland, 23-25 Sept 2020,) pp. 511–515.
[9] I. Boda, M. Bényei, and E. Tóth, "New dimensions of an ancient Library: the Library of Alexandria," in CogInfoCom 2013. Proc. of the 4th IEEE International Conference on Cognitive Infocommunications, (Budapest, Hungary, December 2-5, 2013,) pp. 537–542.
[10] I. Boda, E. Tóth, M. Bényei, and I. Csont, "A three-dimensional virtual library model of the ancient Library of Alexandria," in ICAI 2014. Proc. of the 9th International Conference on Applied Informatics, (Eger, Hungary, January 29-February 1, 2014,) vol. 1, pp. 103–111.
[11] I. Boda, E. Tóth, I. Csont, and L. T. Nagy, "Toward aknowledge base of literary content focusing on the ancient Library of Alexandria in the three dimensional space," in CogInfoCom 2015. Proc. of the 6th IEEE International Conference on Cognitive Infocommunications, (Győr, Hungary, October 19-21, 2015,) pp. 251–258.
[12] I. Boda, E. Tóth, I. Csont, and L. T. Nagy, "Developing a knowledge base of ancient literary texts in virtual space," in P. Baranyi Ed. CogInfoCom 2016. Proc. of the 7th IEEE International Conference on Cognitive Infocommunications, (Wrocław, Poland, October 16-18, 2016,) pp. 263–270.
[13] I. Boda, E. Tóth, I. Csont, and L. T. Nagy, "The use of mythological content in virtual learning environment," in P. Baranyi Ed. CogInfoCom 2017. Proc. of the 8th IEEE International Conference on Cognitive Infocommunications, (Debrecen, Hungary, September 11-14, 2017,) pp. 307–314.
[14] *MaxWhere VR Even more*. [Online]. Available: http://www.maxwhere.com/ (2021-01-05).
[15] I. Boda and E. Tóth, "From Callimachus to the Wikipedia: an ancient method for the representation of knowledge in the WWW era," in Gy. Sallai Ed. CogInfoCom 2018. Proc. of the 9th IEEE International Conference on Cognitive Infocommunications, (Budapest, Hungary, August 22-24, 2018,) pp. 205–210.
[16] I. Boda and E. Tóth, "Text-based approach to second language learning in the virtual space focusing on Callimachus' life and works," in P. Baranyi Ed. CogInfoCom 2019. Proc. of the 10th IEEE International Conference on Cognitive Infocommunications, (Naples, Italy, October 23-25, 2019,) pp. 439–444.
[17] I. Boda and E. Tóth, "English language learning in virtual 3D space by visualizing the library content of ancient texts," in P. Baranyi Ed. CogInfoCom 2020. Proc. of the 11th IEEE International Conference on Cognitive Infocommunications, (Mariehamn, Finland, 23-25 Sept 2020,) pp. 305–311.
[18] R. Gairns and S. Redman, Working words: a guide to teaching and learning vocabulary. Cambridge: Cambridge University Press, cop. 1986. pp. 22–43.
[19] *Oxford Advanced Learner's Dictionary*. [Online]. Available: https://www.oxfordlearnersdictionaries.com/ (2021- 06-15)
[20] C. Tribble and G. Jones, Concordances in the Classroom.A resource book for teachers. Harlow: Longman, 1990.
[21] J. Harmer, The Practice of English Language Teaching. Harlow: Longman, 2009.
[22] M. Hewings, Advanced grammar in use: a self-study reference and practice book for advanced learners of English. Cambridge: Cambridge University Press, cop. 1999. Unit 91. pp. 182–183.
[23] I. Boda, and E. Tóth, "English language learning by visualizing the literary content of a knowledge base in the three-dimensional space," *Ann. Math. Inform.* vol. 53, pp. 45–59, 2021.
[24] J. Scrivener, Learning Teaching. A guidebook for English language teachers. Macmillan Education, 2005. Chapter 11. Lexis. pp. 226-251.
[25] I. Boda and E. Tóth, "Content development for second language learning in the 3D virtual space," in CogInfoCom 2021. Proc. of the 12th IEEE International Conference on Cognitive Infocommunications, (September 23-25, 2021, online on MaxWhere 3D web,) IEEE, 2021. pp. 321–327.
[26] I. Boda, E. Tóth and L. T. Nagy, "Improving A Bilingual Learning Material in the Three-Dimensional Space Using Google Translate," in CogInfoCom 2022. Proc. of the 13th IEEE International Conference on Cognitive Infocommunications, (Online, September 22-23, 2022,) IEEE, 2022. pp. 93–98.
[27] F. E. Weinert, "Introduction and overview: metacognition and motivation as determinants of effective learning and understanding," in F. E. Weinert and R. H. Kluwe, *Metacognition, Motivation and Understanding*. Hillsdale, NJ – London: Lawrence Erlbaum Associates, 1987. pp. 1–16.
[28] G. Tolcsvai Nagy, A magyar nyelv szövegtana [Textology of the Hungarian Language]. Budapest: Nemzeti Tankönyvk.
[29] I. Horváth, "An analysis of personalized learning opportunities in 3D VR," *Front. Comput. Sci.*, 3:673826, 2021.
[30] B. Lampert, A. Pongracz, J. Sipos, A. Vehrerand, I. Horváth, "MaxWhere VR-learning improves effectiveness over classical tools of e-learning," *Acta Polytechnica Hungarica*, vol. 15, no. 3, pp. 125–147, 2018.

István Károly Boda is a college professor at the Debrecen Reformed Theological University, Hungary. He is the head of the Department of Mathematics and Informatics. He was conferred the degree of Ph.D. in 2002; the title of his doctoral dissertation was "Use of Hypertext in Information Science: Concepts, Systems, Models and Applications". In 1994, he was declared a habilitated doctor ("dr. habil.") in the field of mathematics and computer science. His research interest includes computational linguistics and the applications of hypertext especially in text linguistics and stylistics. Presently he investigates the cognitive infocommunication aspects of developing knowledge bases in virtual 2D and 3D environments.

Erzsébet Tóth is an assistant professor at the Faculty of Informatics, University of Debrecen, Hungary. She had a degree in English language and literature and library and information science in 1995. In 2008 she obtained her Ph.D degree in information science and technology. In the doctoral dissertation she investigated the evaluation of search engine performance. She is involved in a virtual library project that focuses on the presentation of digitized library content in three-dimensional space, and she studies the enhanced possibilities of the English language teaching and learning in virtual learning environment. Since 2022 she has been an IEEE member.

Model of the Internet of Digital Education and its links to VR

Attila Kovari

Abstract—**Modern technologies are now essential to make the educational process more mobile, informative and versatile. The digitization of the educational institution, taking advantage of the opportunities provided by the Internet, provides an opportunity to exchange the accumulated experience and knowledge, and provides access to information for everyone. The mandatory introduction of online education due to COVID-19 has accelerated the development of digital education. The main direction in the development of education is moving towards modern online courses, which are already being used successfully. A comprehensive concept is needed to ensure that all IT used for learning, education and training is seamless, free of barriers and closed ecosystems. In connection with the above findings this article first introduces the model and of Internet of Digital Education (IoDE), describes the principles of IoDE, and then summarize the cognitive aspects and VR relations. The purpose of this article is to help readers recognize the IoDE as a whole and predict future trends in IoDE development.**

Index Terms—**Internet of Digital Reality; Digital Transformation; Education; Digital Pedagogy**

I. Introduction

As modern infocommunication technologies become more prevalent, we are moving towards a new digital world characterized by multi-modal entangled combination of IT with human and social cognitive systems [1]. The Internet and several 3D technology as virtual reality (VR), augmented reality (AR), mixed reality together with artificial intelligence (AI), and 5G networks are becoming ubiquitous technologies that affect society as a whole [2]. The impact of modern infocommunication technologies on the digital society is expected to lead to a qualitatively new kind of digital activities [3]. The concept of Internet of Digital Reality (IoD) as a set of network and related technologies for the management, of Digital Realities also have an impact on future forms of life, work and education [1].

As a result of the technologies mentioned above, education is undergoing major changes in recent decades and still is. Students now use different mental models to collect, process, and analyze data [4]. Greater emphasis will be placed on image-based information, scanned reading and responsive online communication [5][6][7]. They also use differen learning styles that include the ability to obtain information from groups and communities within online networking connections.

Attila Kovari is with the Institute of Digital Technology, Faculty of Informatics, Eszterházy Károly Catholic University, Eger, Hungary (e-mail: kovari.attila@uni-eszterhazy.hu) and GAMF Faculty of Engineering and Computer Sciences, John von Neumann University, Kecskemet, Hungary.

Modern technologies are now essential for the educational process to become more mobile, informative and versatile [8]. The digitization of the educational institution, using the opportunities provided by the Internet, provides an opportunity to exchange the accumulated experience and knowledge, and provides access to personalized information for all [9]. The digital transformation of universities means not only the introduction of IT solutions, but also the optimization of existing business processes, which makes them more flexible and “user-friendly” and can give the university greater popularity among students [10].

Presumably, in the near future, artificial intelligence will gain more and more ground in education, paper textbooks will largely disappear and the methodology of education will be transformed [11]. The main direction in the development of education is moving towards online courses, which are already being used successfully [12] [13]. A third major direction of technological change is simulators and virtual reality, which are becoming more widespread and can be used to teach a variety of skills and disciplines [14-20].

A comprehensive concept is needed to ensure that all IT used for learning [21-24], education and training is seamless [25-27], free of barriers and closed ecosystems. In connection with the above findings this article first introduces the importance and describes the principles of IoDE, present the model, then, and then summarize the cognitive aspects and VR relations.

For all IT tools used in learning, education and training to work seamlessly, without barriers and closed systems, a comprehensive approach is needed. The aim of the paper is to help readers to fully understand IoDE and to define the requirements of IoDE.

II. Importance and Principles of IoDE

The IoDE concept is a system for facilitating learning and teaching processes based on the interconnection of the internet and digital education. The term is based on the analogy of the Internet of Things (IoT) [28-33], which refers to the interconnection of various devices and systems via the Internet.

The importance of the concept can be interpreted in several ways:

Access: IoDE enables people around the world to access a variety of digital educational content, facilitating learning and skills development.

DOI: 10.36244/ICJ.2024.5.9

Flexibility: Teaching over the Internet makes the learning process more flexible and personalised, allowing learners to move at their own pace and choose topics that suit their interests.

Resource sharing: the IoDE facilitates the sharing of learning materials, resources and tools between teachers and students, thus increasing the quality and efficiency of education.

Interactive learning: Internet-based digital education gives students the opportunity to learn new skills in an interactive way, for example through virtual labs, games or simulations.

Connected world: the IoDE facilitates global collaboration in education, allowing students and teachers to share experiences and knowledge with each other, no matter where they are in the world.

In summary, the importance of the Internet of Digital Education concept lies in its ability to effective education, increase learning efficiency and enable global collaboration to disseminate knowledge. In the field of IoDE, the principles include efficiency, personalization, and a focus on self-education, which reflect the innovation trends of today's digital education that can be supported by the IoDE concept.

A. *Effectiveness*

It aims to introduce and apply new technological solutions to improve learning. The IoDE concept requires effective packaging and distribution of content between different actors. Effective IoDE based on individual approach to the overall organization of training, ensures interactivity, mobility and flexibility.

B. *Put in context*

Technology provides effective learning when we put things in context and it is led by a teacher. The use of technology needs to be placed in a context that meets the goals of education, we need to know how to find the answers using technology.

C. *Consistency*

An effective IoDE must offer a training ecosystem that provides a system of educational resources that combines the needs of students, universities, and companies interested in training so that the relationship between training opportunities is clear and traceable.

D. *Internet and AI*

The data processing capability of artificial intelligence and the connections provided by the Internet can bring many benefits to learning. For example, if as much data as possible is available, we can learn more from analyzing the information obtained from the analysis of that data. AI algorithms can extract and compare data from different learning environments, for example, to show which activities provide the most effective learning, the best results. These may include learning management systems, interactive learning environments, or even educational games, but not necessarily educational games that can enable data-rich learning through play. The importance of the information provided by the data is based on the fact that it can be used to plan and implement a personalized learning and teaching process that can optimize the level of competence of students with their learning activities.

E. *Personalization*

IoDE systems must ensure that the principle of one-size-fits-all content prevails over the principle of equal content for all learners, which is typical of the traditional education system. Adaptive learning systems driven by artificial intelligence provide an opportunity to consider formal and informal learning opportunities in addition to students' level of knowledge. The intelligent learning system adapts the teaching process and teaching materials to the individual learning speed of the students, highlighting the key concepts of the given curriculum and encouraging the learner to focus on the content of the curriculum that has not yet been mastered.

F. *High level of self and AI assisted learning*

The application of artificial intelligence in online learning has recently become an important new issue and method. Learning driven by artificial intelligence is of paramount importance in supporting self-directed learning, and this is partly consistent with teaching a human through machine control.

G. *Interoperability*

There is a need for standards that include requirements for providers of educational information technology programs that can ensure that technology is independent of devices and ecosystems and is fully interoperable [38]. This avoids that learners can only complete the learning activity in a given IoDE system, and effective learning is easily interoperable between systems, which is a prerequisite for the effective implementation of the IoDE concept. Some of the more significant standards related to learning are summarized in Figure 1.

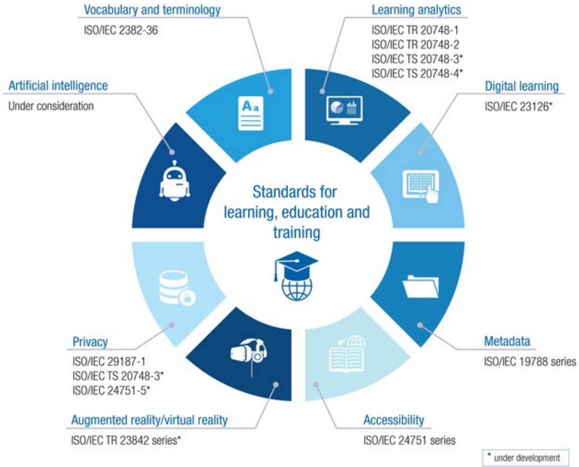


Fig. 1. Standards for learning, education and training [38].

H. Data security

In addition, data security and privacy must be guarantee for digital learning.

III. Generalized Model of IoDE

IoDE aims to provide a comprehensive approach to making teaching and learning processes more efficient and effective in the digital age, using the opportunities offered by the internet [34-36]. From the perspective of the IoDE model, an environment can be created for learners and educators to achieve personalized teaching and learning, taking into account individual goals, by interweaving the Internet and digital education. Based on the above, the IoDE model is shown in Fig. 2.

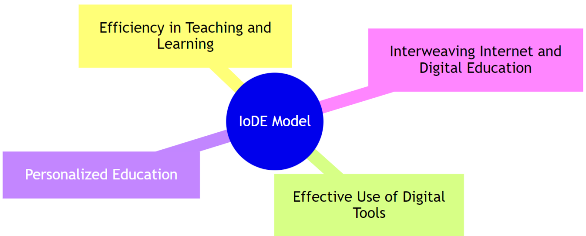


Fig. 2. Generalized model of IoDE

The model must reflect the methodologies shown in Figure 2, with particular regard to generations and learning styles. or early learners, IoDE can incorporate interactive and visually engaging content that aligns with their developmental stage. This includes intuitive interfaces, gamification, and multimedia resources that cater to their limited attention span and developing cognitive abilities. In case of school-age children, IoDE can offer more structured learning paths with a mix of interactive activities and educational content. Customization to accommodate varying learning speeds and styles is crucial, as is the integration of collaborative tools for peer interaction. For teenagers, IoDE can focus on more advanced, subject-specific content. This age group benefits from a blend of autonomy in learning and guided pathways. Tools for self-assessment, forums for discussion, and resources for project-based learning are particularly effective. Adult education through IoDE should focus on flexibility and relevance. This includes modular courses, career-oriented content, and resources that accommodate various life commitments. Adult learning often emphasizes practical, skill-based education that can be directly applied in the workplace.

IoDE can offer comprehensive curriculum and resources for parents and students engaged in home schooling. It provides access to diverse educational materials, interactive tools, and community forums for collaboration and support. In vocational training, IoDE can be instrumental in providing practical, skills-based learning through simulations, instructional videos, and interactive modules, catering to specific industry needs. IoDE enhances online education by offering a broad range of courses and resources accessible from anywhere. This includes MOOCs (Massive Open Online Courses), webinars, and virtual classrooms. IoDE plays a crucial role in making education more accessible and inclusive. It breaks down geographical and physical barriers, provides resources for learners with disabilities, and offers content in multiple languages to cater to a global audience.

As summarized above, a proposed detailed model for the Internet of Digital Education (IoDE) is built around the main elements and factors listed in Figure 3 and details are summarized in Table I.

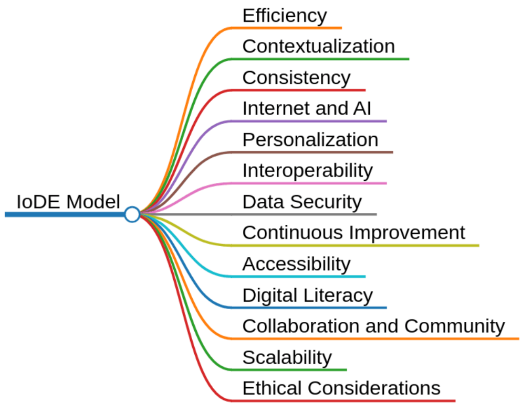


Fig. 3. Main factors of IoDE

TABLE I
Main elements and factors of IoDE

Factors	Details
Efficiency	Implementing and using new technologies to improve learning
	Use of new technologies to introduce new technologies and technologies for the delivery of training
Contextualization	Using technology in a context that is consistent with educational objectives
	Empowering teachers to help students find answers through technology
Consistency	Creating a training ecosystem that combines the needs of students, universities, and interested companies
	Clarity and traceability of links between training opportunities
Internet and AI	Leveraging the computing capabilities and benefits of AI and internet connectivity for learning
	Using AI algorithms to extract and compare data from different learning environments
	Using data to design and implement personalized learning and teaching processes
Personalization	Preference for personalized, universal content over content that is the same for all learners
	AI-driven adaptive learning systems that take into account formal and informal learning opportunities
	Adapting the learning process and teaching materials to individual learning speeds
	High levels of autonomous and AI-assisted learning
	AI-driven learning is paramount in supporting autonomous learning, partly in line with machine-driven instruction
Interoperability	Establishing prescriptive standards for education IT software providers

	Easy and efficient interoperability of effective learning between systems
Data Security	Ensuring data security and privacy for digital learning participants
Continuous Improvement	Regular updates and refinements of IoDE systems based on user feedback and technological developments
	Encouraging collaboration between educators, researchers, and technology providers
Accessibility	Ensuring IoDE systems are accessible to learners of different abilities, backgrounds, and learning preferences
	Application of inclusive design principles to promote equitable learning experiences
Digital Literacy	Teaching digital literacy skills to students, educators, and administrators
	Developing training programs and resources to support the development of digital literacy
	Incorporating AI-driven assessment and evaluation tools to measure student progress and learning outcomes
	Using data-driven insights to inform instructional design and curriculum development
Collaboration and Community	Building a sense of community between learners and educators in the IoDE environment
	Using technology to facilitate collaboration, communication, and peer support
Scalability	Designing IoDE systems to be scalable and adaptable to meet growing user numbers and changing needs
	Encouraging the use of IoDE solutions in different educational environments
Ethical Considerations	Addressing ethical issues related to AI in education, such as bias, fairness, and transparency
	Developing guidelines and rules to ensure responsible AI integration in the IoDE ecosystem

The IoDE model and concept brings several innovations to education compared to other concepts. IoDE develops education using digital technologies and the Internet, allowing learners and teachers to access educational materials and resources quickly and easily. Focusing education on collaboration. The IoDE concept can indirectly contribute, through the opportunities provided by the Internet, to enabling learners to progress adaptively at their own pace under flexible conditions. Taking advantage of global teaching-learning opportunities can increase efficiency [37-39].

IV. Cognitive Aspects of IoDE

The cognitive aspects of the Internet of Digital Education (IoDE) relate to the ways of learning, cognitive development and knowledge acquisition supported by technology and AI-driven solutions. The main cognitive aspects are:

A. Personalization

IoDE systems adapt to individual learning speeds, preferences and knowledge levels, ensuring a personalised learning experience.

B. Contextualization:

Technology is applied in a context that is consistent with educational goals, helping learners make meaningful connections and better understand complex concepts.

C. Interactivity

IoDE promotes active participation, allowing learners to interact with content and collaborate with peers and instructors, resulting in deeper understanding and improved problem-solving skills.

D. Independent learning

AI-assisted learning supports independent learning, helping students to achieve their own goals and define their own development path.

The importance of the cognitive aspects in the context of the Internet of Digital Education (IoDE) can be attributed to a number of factors. A personalized learning experience facilitates the individual development and progress of learners, taking into account their different abilities, interests and prior knowledge. A contextualization helps learners to understand and apply newly acquired knowledge to real-life situations, increasing the effectiveness of learning. Interactivity and active participation foster learners' critical thinking, problem-solving skills and deeper understanding of concepts. The independent learning promotes learners' self-activity and responsibility for their own learning process, which has a positive impact on motivation and achievement in the long term.

Taking cognitive aspects into account allows learning processes to be made more flexible to better adapt to learners' changing needs and circumstances. Equal access and equal opportunities: taking into account the cognitive aspects of IoDE promotes equal access and equal opportunities in learning by providing all learners with the opportunity to benefit from an optimal learning environment.

Overall, the importance of the cognitive aspects lies in their ability to enable more effective, personalised and flexible learning experiences, facilitating learners' success and progress in IoDE systems.

V. VR in the Concept of IoDE

Virtual Reality can be used in the context of the cognitive aspects of IoDE as a tool to further enhance learning effectiveness, provide students with a deeper understanding and personalized learning experiences that support individual development. The linking of virtual reality (VR) to the concept of IoDE mainly focusing on the following aspects:

- Immersive learning experiences: VR allows learners to be fully immersed in a virtual environment that provides rich, interactive and realistic experiences during the learning process.
- Practical application: VR allows learners to try out and practise newly acquired knowledge in a safe and controlled environment, contributing to a more effective acquisition of skills and competences.
- Customizability and adaptability: VR environments are easily customizable and adapt to learners' individual needs, prior knowledge and learning pace, facilitating a personalized learning experience.

- Remote collaboration: VR enables learners and educators to collaborate and interact with each other despite any geographical distance, facilitating global collaboration and communication.
- Facilitated contextualisation: VR can be particularly useful in presenting complex or difficult-to-access concepts, phenomena and situations, helping learners to understand the context and apply knowledge to real-life situations.
- Maintaining motivation and interest: the exciting and interesting learning experiences provided by VR can increase learners' motivation and interest in the subject matter.

By combining VR and IoDE, it creates an immersive and interactive learning environment for both students and teachers, which significantly enhances the learning experience and the effectiveness of teaching. The introduction of VR technology on IoDE platforms allows for more visualisation of learning materials and increased learner engagement. The use of virtual reality in education facilitates distance learning and global collaboration, allowing students to access learning materials and tutor support anywhere and anytime. Combining VR and IoDE in education also contributes to supporting students with special needs. Instruction in virtual reality helps to better serve different learning styles and abilities, as well as offer learning experiences tailored to the individual needs of students. The combination of IoDE and VR allows for the continuous development and innovation of educational processes using new methods and tools. For educators, teaching in virtual reality opens up new opportunities for teaching and learning that more effectively encourage students to be creative, collaborative and problem-solving. Therefore, the link between virtual reality and the Internet of Digital Education can make a significant contribution to improving the learning experience, teaching methods and individual development in the field of education.

VI. Requirements for implementation of IoDE

The challenges of the 21st century require creative and innovative thinking and a changing division of labor between workers and machines as a result of the spread of artificial intelligence [39]. Digital competencies are gaining in value, bringing to the fore the knowledge, skills and attitudes that people can use to achieve their goals effectively [40][41]. Digital skills focus on the technological, communication, information and multimedia aspects that lead to consciously effective action to create value [42] [43].

Important changes are taking place in the areas of intelligent manufacturing, fully virtualized interfaces, and remote control devices. Future professionals need to have different competencies than before, so they need to be developed. The increasingly complex interaction of people with modern technologies will bring significant changes. The field of cognitive infocommunication [44] will become even more important, in which digital competencies problem-solving, creativity and collaboration [45-48], effective communication [49] and study results [50], emotional intelligence [51-53], BigData [54], 3D [55], VR [56] or effective use of modern digital devices [57-59]. In addition, human-computer interfaces (HCIs) such as eye-tracking [60-62] or brain-computer interfaces (BCIs) [63] can be an alternative possibility to analyze the effectiveness of different cognitive processes.

The need to combine advanced technological and soft skills requires the use of educational environments and methodologies that are based on the integration of multiple disciplines, allowing the emergence of IoDE applications. The application of IoDE can effectively improve digital development of competencies in the Internet, learning by integrating spaces and educational experiences that can lead to effective competence development.

The most basic digital competencies of current learners have already developed to a greater extent, as they have grown up in the digital environment since childhood. Virtual online learning environments are already available to students and many technological opportunities are not yet fully exploited, we are witnessing an incredible transformation of educational environments that lay the foundations for digital educational institutions and digital universities. The days, activities, communication and interactions of a significant part of the students and already the teachers with the outside world take place largely through the Internet. Learning styles are also changing and changing, with students now using different methods of data collection and processing than they did 10-20 years ago. Image and video-oriented information, scanned reading, and information from multiple sources that have developed a different learning style in students come to the fore. These all predict the effective use of IoDE.

The development of ICT enables the introduction of many innovative teaching techniques and methodologies in distance learning, also in the case of flipped classroom or hybrid learning techniques.

It can also be seen from the above that an essential condition for effective IoDE is that the technology is available at the right level, in the right number and with the right support in the educational institution. Of course, this is not available in all institutions and at the same level, not only in terms of digital technology tools, but also in terms of the knowledge, methodology, useful experience, ideas and elaborated curricula needed to use them. Understanding of digitization and its transformative effects, developing and integrating a digital switchover strategy, and exploiting the potential of IoDE are key factors in implementing effective IoDE.

In summary, the main requirements for effective IoDE implementation are:

- in general, the institution should be aware of digitization and its transformative effects, understand the role of digitization and provide appropriate support;

- the institution should plan ahead, its application should be an important goal in its vision and strategic plan, the staff and faculty of the institution should be involved and support the transformation, and proper communication is essential;

- the institution should support those at the forefront of transformation, help and support the testing of new and innovative working methods, including the provision of IT professional support;

- take advantage of the systemic and community-level opportunities provided by IoDE to increase the effectiveness of education and to adapt to modern forms of learning;

- students and staff should be kept up to date with the technological and methodological potential of IoDE.

VII. Conclusion

The nature of education known today is facing a significant change due to the emergence of advanced technologies. Higher education institutions compete for students, play an important role in this competition, and have the potential to benefit from advanced digital communication technologies based on the Internet and methodologies based on them in education.

University leaders need to recognize that there is an increasing need to modernize their education systems and processes, and to transform their pedagogical, teaching methodology, and curriculum, subject, and thematic practices. Institutions develop unique plans and strategies, try to introduce unique digital solutions in order to take advantage of the opportunities provided by advanced technologies, but the most effective solutions do not necessarily come to the fore during the implementation. Resources devoted to the development of ICT systems do not always bring the desired results. The present study approaches the exploitation of the Internet-based potential of digital education through the conceptual definition of the digital framework introduced with the introduction of IoDE.

IoDE ecosystem connects students, teachers, and other cognitive entities, digital education systems using the combinations of most relevant technologies and data that create a higher-level functional integration to achieve higher-level learning efficiency. This could be one of the foundations for the successful educational institutions of the future.

References

[1] P. Baranyi, Á. Csapó, T. Budai, Gy. Wersényi, “Introducing the Concept of Internet of Digital Reality – Part I,” *Acta Polytechnica Hungarica*, vol. 18, no. 7, pp. 225–240, 2021, **DOI**: 10.12700/APH.18.7.2021.7.12.

[2] J. Orlosky, K. Kiyoshi, T. Haruo, “Virtual and augmented reality on the 5G highway,” *Journal of Information Processing*, volt 25, pp. 133–141, 2017. **DOI**: 10.2197/ipsjjip.25.133

[3] T. Lengyel Molnár, “Library strategies in the Cyber-Physical Society,” *Proceedings of Society for Information Technology & Teacher Education International Conference 2020*, Waynesville (NC), Association for the Advancement of Computing in Education, pp. 1048–1054, 2020.

[4] K. Chmielewska, “From Mathability to Learnability”, *Proceedings of the 10th IEEE International Conference on Cognitive Infocommunications*, 2019, pp. 513–516. **DOI**: 10.1109/CogInfoCom47531.2019.9089891

[5] K. Chmielewska, “Towards‘ Learnability’ – What we benefit from mathability”, *Acta Polytechnica Hungarica*, vol. 19, no. 1, pp. 171–192, 2022. **DOI**: 10.12700/APH.19.1.2022.19.11

[6] K. Chmielewska, A. Gilányi, “Educational context of mathability”, *Acta Polytechnica Hungarica*, vol. 15, no. 5, pp. 223–237, 2018. **DOI**: 10.12700/APH.15.5.2018.5.13

[7] K. Chmielewska, A. Gilányi, A. Łukasiewicz, “Mathability and mathematical cognition”, Proceedings of the 7th IEEE Conference on Cognitive Infocommunications, 2016, pp. 245–250. **DOI**: 10.1109/CogInfoCom.2016.7804556

[8] R. Racsko, E.M. Troll. “Our digital education habits in the light of their environmental impact: the role of green computing in education,” *Teaching Mathematics And Computer Science*, vol 18 no. 1, pp. 69–86, 2020.

[9] I. Horváth, “An Analysis of Personalized Learning Opportunities in 3D VR,” *Frontiers Computer Science*, vol 20, ArtNo: 673826, 2021. **DOI**: 10.3389/fcomp.2021.673826

[10] M. Rajcsányi-Molnár, S. Czifra, “Development of an integrated online education system at the University of Dunaújváros,” *Fenntarthatósági terek*, DUE Press, 2021, pp. 84–101.

[11] P. Négyesi, I. Oláh Téglási, R. Racsko, “The application of a custom-made mathematical software in order to support problem-solving strategy of high school students,” *Proceedings of XXXIII. DidMatTech 2020 Conference, New Methods and Technologies in Education in Research and Practice*, Budapest, Hungary, 2020, pp. 402–407.

[12] B. Orosz, C. Kovács, D. Karuovic, Gy. Molnár, L. Major, V. Vass, Z. Szűts, Zs. Námesztovszki, “Digital education in digital cooperative environments,” *Journal of Applied Technical and Educational Sciences*, vol 9, no 4, pp. 55–69, 2019. **DOI**: 10.24368/jates.v9i4.149

[13] K. Chmielewska, W. Ciskowska, D. Glazik, D Marcinek, K. Wojciechowska, A. Gilányi, “Learnability – are we ready for distance learning?”, *Proceedings of the 11th IEEE International Conference on Cognitive Infocommunications*, 2020, pp. 465–470. **DOI**: 10.1109/CogInfoCom50765.2020.9237881

[14] M. Dergham, A. Gilányi. “Application of virtual reality in kinematics education”, *10th IEEE International Conference on Cognitive Infocommunications*, 2019, pp. 107–111. **DOI**: 10.1109/CogInfoCom47531.2019.9089971

[15] A. Gilányi, A. Rácz, A. M. Bólya, K. Chmielewska, “Early history of Hungarian ballet in virtual reality”, *10th IEEE International Conference on Cognitive Infocommunications*, 2019, pp. 193–198. **DOI**: 10.1109/CogInfoCom47531.2019.9089943

[16] J. Katona, “A Review of Human–Computer Interaction and Virtual Reality Research Fields in Cognitive InfoCommunications,” *Applied Sciences*, vol. 11, no. 6, pp. 2646, 2021. **DOI**: 10.3390/app11062646

[17] A. Kovari: Possibilities of applying project education in IT education. In *XXVI. International Conference on Computer Science and Education*, 2016, pp. 247–251.

[18] M. Dergham, A. Gilányi, “On a system of virtual space for teaching kinematics”, *10th IEEE International Conference on Cognitive Infocommunications*, 2019, pp. 411–414. **DOI**: 10.1109/CogInfoCom47531.2019.9089935

[19] A. Gilányi, A. Rácz, A. M. Bólya, J. Décsei, K. Chmielewska, “A presentation room in the virtual building of the first National Theater of Hungary”, *11th IEEE International Conference on Cognitive Infocommunications*, 2020, pp. 517–522. **DOI**: 10.1109/CogInfoCom50765.2020.9237828

[20] A. Rácz, A. Gilányi, A. M. Bólya, J. Décsei, K. Chmielewska, “On a model of the first National Theater of Hungary in MaxWhere”, *11th IEEE International Conference on Cognitive Infocommunications*, 2020, pp. 573–574. **DOI**: 10.1109/CogInfoCom50765.2020.9237848

[21] B. Holmberg, *Theory and Practice of Distance Education, Routledge*, 2005.

[22] Gy. Molnár, “Teaching and Learning in modern digital Environment,” *Proceedings of the IEEE 13th International Symposium on Applied Machine Intelligence and Informatics*, pp. 213–217. **DOI**: 10.1109/SAMI.2015.7061878

[23] E. Gőgh, R. Racsko, A. Kővári, "Experience of Self-Efficacy Learning among Vocational Secondary School Students," *Acta Polytechnica Hungarica*, vol. 18 no. 1, pp. 101–119, 2021. **DOI**: 10.12700/APH.18.1.2021.1.7

[24] T. Lengyelné Molnár, "Changing reading habits and methodological options resulting from digital transformation," *Journal of Applied Technical And Educational Sciences*, vol. 9, no. 4, pp. 27–42, 2019. **DOI**: 10.24368/jates.v9i4.108

[25] M. Rajcsányi-Molnár, A. Bacsa-Bán, "From the Initial Steps to the Concept of Online Education" – Teacher Experiences and Development Directions Based on Feedback from Online Education Introduced During the Pandemic. *Central European Journal of Educational Research*, vol. 3, no 3, pp. 33–48, 2021. **DOI**: 10.37441/cejer/2021/3/3/9612

[26] Gy. Molnár, Zs. Namestovski, R. Pinter, "Massive Open Courses (MOOCs) held in Serbian," *Journal for Information Technology Education Development and Teaching Methods of Technical and Natural Sciences*, vol 6, no 1, pp. 12–16, 2016.

[27] D. E. Salinas-Navarro, C. L. Garay-Rondero, "Experiential learning in Industrial Engineering education for Digital Transformation," *2019 IEEE International Conference on Engineering, Technology and Education (TALE)*, 2019, pp. 1–9, **DOI**: 10.1109/TALE48000.2019.9225984.

[28] K. Thorsten, R. Van Kranenburg, S. Lange, "Introduction to the Internet of Things," *Enabling Things to Talk*. Springer, Berlin, Heidelberg, 2013, pp. 1–10. **DOI**: 10.1007/978-3-642-40403-0_1

[29] L. Srivastava, T. Kelly, "The internet of things," International Telecommunication Union, Tech. Rep 7, 2005.

[30] T. Guzsvinecz, T. Medvegy, V. Szucs, "A brief review on the challenges of Internet of Things and their solutions," *Proceedings of the Pannonian Conference on Advances in Information Technology (PCIT'2019)*, 2019, pp. 181–186.

[31] E. Harald, P. Goodridge, J. Haskel, "The Internet of Things and economic growth in a panel of countries," *Economics of Innovation and New Technology*, vol 30, no 3, pp. 262–283, 2021. **DOI**: 10.1080/10438599.2019.1695941

[32] Cisco, "The Internet of Everything - Cisco IoE Value Index Study." 2013. https://www.cisco.com/c/dam/en_us/about/business-insights/docs/ioe-value-index-faq.pdf

[33] Gy. Wersényi, Á. Csapó, T. Budai, P. Baranyi, "Internet of Digital Reality: Infrastructural Background – Part II," *Acta Polytechnica Hungarica*, vol. 18, no. 7, pp. 225–240, 2021, **DOI**: 10.12700/APH.18.8.2021.8.5.

[34] J. Lengsfeld, *Digital Era Framework*, 2019. https://joernlengsfeld.com/files/digital-era-framework-en.pdf

[35] Z. Balogh, C. Klimeš, "Modelling of education process in LMS using Petri nets structure," *Proceedings of the IADIS International Conference e-Learning, MCCSIS*, Freiburg, 2010, pp. 289–191.

[36] P. Porteleki, "IoT integration in Microsoft Dynamics ERP," Computers & Learning, vol 3, no 1, pp. 1–11, 2020.

[37] S. Oza *et al*., "IoT: The Future for Quality of Services," In: ICCCE 2019. Lecture Notes in Electrical Engineering, vol 570. Springer, Singapore, 2020, **DOI**: 10.1007/978-981-13-8715-9_35
https://en.wikipedia.org/wiki/Fourth_Industrial_Revolution#/media/File:Internet_of_Things_(IoT).png

[38] A. Price, "Why the future of learning is digital, Digital technologies are transforming learning and building a more sustainable future," e-tech, https://medium.com/e-tech/advancing-the-digitization-of-education-8e47362b2165

[39] World Economic Forum, "The future of jobs report 2018," Geneve: WEF, 2018. https://www3.weforum.org/docs/WEF_Future_of_Jobs_2018.pdf

[40] Gy. Molnár, D. Sik, "The virtual toolkit of digital instruction and its application in digital work forms," *Proceedings of the 2020 11th IEEE International Conference on Cognitive Infocommunications*, 2020, pp. 599–602. **DOI**: 10.1109/CogInfoCom50765.2020.9237855

[41] Gy. Molnár, Z. Szűts, Z. Balogh, "Modern digital Web 2.0 devices and services supporting the teaching of Technology and Informatics," *Proceedings of the IEEE 17th International Symposium on Intelligent Systems and Informatics Proceedings*, 2019, pp. 89–94. **DOI**: 10.1109/SISY47553.2019.9111555

[42] T. Guzsvinecz, J. Szűcs, "Using Analytics to Identify When Course Materials Are Accessed Relative to Online Exams during Digital Education", *Education Sciences*, vol. 11 no. 10, p. 576, 2021. **DOI**: 10.3390/educsci11100576

[43] Gy. Molnár, K. Nagy, Z. Balogh, "The role and impact of visualization during the processing of educational materials, presentation options in education and in the virtual space," *Proceedings of the 2019 10th IEEE International Conference on Cognitive Infocommunications*, 2019, pp. 533–537. **DOI**: 10.1109/CogInfoCom47531.2019.9089917

[44] D. Mattyasovszky-Philipp, B. Molnár, "An Architectural Approach to Cognitive Information Systems," *Acta Polytechnica Hungarica*, vol 17, no 2, pp. 237–253, 2020. **DOI**: 10.12700/APH.17.2.2020.2.13

[45] G. V. Kövecses, B. Lampert, A. Pongrácz, and I. Lőrincz, "University Lecturers' Distance Learning Experiences Gained during the COVID-19 Pandemic Period," *Proceedings of the 2020 11th IEEE International Conference on Cognitive Infocommunications*, 2020, pp. 325–330. **DOI**: 10.1109/CogInfoCom50765.2020.9237890

[46] T. Guzsvinecz, É. Orbán-Mihálykó, E. Perge, and C. Sik-Lányi, "Analyzing the Spatial Skills of University Students with a Virtual Reality Application using a Desktop Display and the Gear VR," *Acta Polytechnica Hungarica*, vol. 17, no. 2, pp. 35–56, 2020. **DOI**: 10.12700/APH.17.2.2020.2.3

[47] Z. R. Rajtik, "Vocational training 4.0 strategy ideas for the field of mechanical engineering," *Transactions on IT and Engineering Education*, vol 3, no 1, pp. 16–31, 2020.

[48] M. Rajcsányi-Molnár, "MaTech: Organizing a creative math competition based on digital tools for secondary school students," *Felsőoktatási innovációk a tanulás korában*, MELLearN, 2019, pp. 19–32.

[49] D. Mattyasovszky-Philipp, A. M. Putnoki, B. Molnár, "The Unrepeatable Human Mind – Challenges in the Development of Cognitive Information Systems – What Makes a Machine Human?," *Electronics* 2022, vol 11, ArtNo 394. **DOI**: 10.3390/electronics11030394

[50] Z. Balogh, M. Kuchárik, "Predicting Student Grades Based on Their Usage of LMS Moodle Using Petri Nets," *Applied Sciences,* vol 9, no 20, ArtNo: 4211, 2019. **DOI**: 10.3390/app9204211

[51] G. V. Kövecsesné, A. Benyák, A. Makkos, and Á. Fehér, "Emotional intelligence development in online education," *Proceedings of the 12th IEEE International Conference on Cognitive Infocommunications*, 2021, pp. 517–521.

[52] M. Magdin, *et al*., "Automatic detection and classification of emotional states in virtual reality and standard environments (LCD): comparing valence and arousal of induced emotions," *Virtual Reality,* vol 25, no 4, pp. 1029–1041, 2021. **DOI**: 10.1007/s10055-021-00506-5

[53] A. Makkos, Á. Fehér, A. Mészáros, G. V. Kövecsesné, A. Benyák, "Examining the emotional intelligence and coping strategies of trainee teachers and in-service teachers," *ICFE 2021, School and teacher education* 2030, 2021, p 18.

[54] A. Benedek, Gy. Molnár, Z. Szűts, "Practices of Crowdsourcing in relation to Big Data Analysis and Education Methods," *Proceedings of the IEEE 13th International Symposium on Intelligent Systems and Informatics*, 2015, pp. 167–172. **DOI**: 10.1109/SISY.2015.7325373

[55] Z. Szűts, "An Iconic Turn in Art History – The Quest for Realistic and 3D visual Representation on the World Wide Web," *The Iconic Turn in Education (Visual Learning)*, Frankfurt am Main, Peter Lang, 2012, pp. 59–66.

[56] I. Horváth, "Innovative engineering education in the cooperative VR environment," *Proceedings of the 7th IEEE International Conference on Cognitive Infocommunications*, 2016. pp. 359–364. **DOI**: 10.1109/CogInfoCom.2016.7804576

[57] G. V. Kövecsesné, "Supporting the learning of teaching students with digital tools," *Journal of Applied Technical and Educational Sciences*, vol. 10, no. 4, pp. 105–124, 2020. **DOI**: 10.24368/jates.v10i4.208

[58] V. Szücs and C. Steiner, "Multifunctional Chatbot for search support and secure the host," *Proceedings of the 12th IEEE International Conference on Cognitive Infocommunications*, 2021, pp. 943–948.

[59] I. Horváth, "MaxWhere 3D Capabilities Contributing to the Enhanced Efficiency of the Trello 2D Management Software", *Acta Polytechnica Hungarica*, Vol. 16, No. 6, pp. 55–71. 2019. **DOI**: 10.12700/APH.16.6.2019.6.5

[60] J. Katona, "Measuring Cognition Load Using Eye-Tracking Parameters Based on Algorithm Description Tools," *Sensors*, vol. 22, no. 3, pp. 912, Jan. 2022. DOI: 10.3390/s22030912

[61] J. Katona, "Analyse the Readability of LINQ Code using an Eye-Tracking-based Evaluation," *Acta Polytechnica Hungarica*, vol. 18, no. 1, pp. 193–215, 2021. DOI: 10.12700/APH.18.1.2021.1.12

[62] J. Katona, "Clean and Dirty Code Comprehension by Eye-tracking Based Evaluation using GP3 Eye Tracker," *Acta Polytechnica Hungarica*, vol. 18, no. 1, pp. 79–99, 2021. DOI: 10.12700/APH.18.1.2021.1.6

[63] J. Katona, "Examination and comparison of the EEG based attention test with CPT and TOVA," in *15th IEEE International Symposium on Computational Intelligence and Informatics*, Budapest, Hungary, 2014, pp. 117–120. DOI: 10.1109/CINTI.2014.7028659

Attila Kovari is Professor at Eszterhazy Karoly Catholic University, University of Dunaujvaros, Obuda University and John von Neumann University, Hungary. He holds a Ph.D. degree in Engineering and Habilitation in Computer Engineering and Education. His research areas are human-computer interaction, brain-computer-interfaces, eye-tracking analysis, IT in industrial applications.

Special Issue of the Infocommunication Journal

Textual Analysis of Virtual Reality Game Reviews

Tibor Guzsvinecz, and Judit Szűcs

Abstract—**Virtual reality systems are complex and made of various parts. Since a person is an integral component of such systems, virtual reality technologies also have a cognitive aspect. As such, these technologies engage with the perceptual, attentional, and decision-making processes of users. Consequently, they can be considered cognitive tools. Thus, it is imperative to understand what people think of such environments. To take a step towards this understanding, textual reviews of virtual reality games made for entertainment were investigated using text mining methods. Thus, 1,635,919 textual reviews were scraped from the Steam digital video game distribution platform in the spring of 2023. The reviews were grouped by whether they were positive or negative. According to the results, the following conclusions can be made regarding virtual reality games: 1) Negative reviews are significantly longer than positive ones. 2) Negative reviews are written significantly earlier than positive ones, although no correlation was found between the review type and the playtime before writing the review. 3) The most frequent words and word correlations are different between review types since negative reviews are more focused on game mechanics and bugs. Due to the results, insights can be provided to virtual reality game developers to help them refine their games.**

Index Terms—**computer games; game review; player experience; Steam; textual analysis**

I. Introduction

In this digital age, the popularity of virtual reality (VR) applications is unquestionable. As VR is a synthetic reality, it is possible to create virtual environments (VEs) that are either similar to or different from the real world. Depending on the goal of such virtual spaces, they could be created for various purposes [1, 2]. For example, VR technologies play an important role in education [3–6], training [7, 8], healthcare [9, 10], and even entertainment [11, 12].

As can be expected based on these previously mentioned fields, a VR system requires a person to interact with it to work perfectly [13]. This is due to the fact that these systems have cognitive aspects as well [14, 15]. VR can also be considered a cognitive tool [16]. Therefore, when designing such systems, it is important to keep the target groups in mind during the process. Naturally, after a VR system or application is implemented, users can experience it and leave feedback about its strengths and weaknesses. This is also the case with VR applications for research or entertainment purposes. Feedback can also be verbal or textual.

Understanding textual feedback is crucial to know how people reacted to various experiences. This textual feedback can come in the form of a review [17, 18], which allows game developers to learn how to improve their games [19]. Video game reviews can contain several topics such as achievements, accessories, general experience, social interaction, social influence, narrative, visual/value, and information about bugs in the game [20]. Thus, it is possible to know how playable a game is from the reviews [21]. Reviews can also detail several factors that can make a game popular [22]. They can also serve as a product review: they can influence other players whether to buy or not to buy a game [23, 24]. Also, critics tend to highlight different aspects of games than players [25].

As can be seen, reviews contain critical information about the experiences of players. Furthermore, many reviews are written daily, thus it is possible to analyze a large number of data [26]. Due to this data, game developers and researchers can ascertain the strengths and weaknesses of the analyzed games. Naturally, these can improve the quality of future games [27, 28].

Thus, the goal of this paper is to take a step towards understanding VR games and applications made for entertainment using textual analysis methods. For this, the following research questions (RQs) were formed:

- RQ1: How long are the reviews of VR games?
- RQ2: Is there a correlation between playtime and the length of VR game reviews?
- RQ3: What are the most frequent words and their word associations in the reviews?

The first question aims to investigate the length of textual reviews for VR games, comparing word count of positive and negative reviews. The second one seeks to explore whether there is a relationship between playtime and the length of reviews. By examining potential correlations, it is possible to determine if users who spend more time playing VR games are more likely to write longer reviews, and vice versa. Lastly, by answering the third one we can uncover key terms, expressions, or themes that emerge frequently in both positive and negative reviews. Thus, understanding reviews can provide insights into the level of detail and depth of user feedback, which can be valuable for developers and researchers.

Submitted on...

Tibor Guzsvinecz is with the Department of Information Technology and its Applications, Faculty of Information Technology, University of Pannonia, Zalaegerszeg, Hungary (e-mail: guzsvinecz.tibor@zek.uni-pannon.hu). He is the corresponding author.

Judit Szűcs is with the Department of Information Technology and its Applications, Faculty of Information Technology, University of Pannonia, Zalaegerszeg, Hungary (e-mail: szucs.judit@zek.uni-pannon.hu)

DOI: 10.36244/ICJ.2024.5.10

Therefore, this paper is structured as follows. Section 2 details the Steam scraping process and the textual analysis. The results are presented in section 3. In section 4, the results are discussed along with the limitations of the study. Lastly, conclusions are made in section 5.

II. Materials and methods

The reviews on the Steam digital game distribution platform were chosen for the analysis. The platform was chosen because it was the largest digital game distribution platform in 2017 [29]. In 2021, it was still one of the biggest platforms with more than 50,000 video games in its library [30]. Not to mention, due to the Covid-19 pandemic, digital video game distribution platforms gained popularity as well [31].

On Steam, players have the opportunity to write reviews on a game's page, although they have to register a free profile and play the said game beforehand. Naturally, it is free to browse and read reviews on the platform. Valve Corporation – the creator of this platform – has developed an application programming interface (API) to freely access and scrape some of the platform's contents, including video game reviews [32].

Thus, this section presents the following. In the first subsection, the reviews on Steam are defined. Afterward, the scraping process is detailed in the second subsection. Lastly, the textual analysis is shown in the third subsection.

A. Video game reviews on Steam

On the Steam digital video game distribution platform, each game has its own store page. These pages contain the reviews as well. According to the Steam API, each review object has multiple components that can be accessed [33]. The following components were used for this research: the textual part, the language of the review, playtime when writing the review, and whether a review is positive or not. For the latter, it should be mentioned that Steam does not have a numerical rating system. Thus, players can either recommend the game or not. The first represents a positive review, while the latter is a negative one. Naturally, in the previously mentioned textual component, players can write down their experiences with no character limitation.

B. The scraping process

The statistical program package R was used for the scraping process along with its *rvest* and *httr* packages. The Steam API was also used to access content during the process. The scraping process was conducted in the spring of 2023.

First, the list of games and applications was created using the *GetAppList* function that could be found in the Steam API. It returned the list of all application IDs on the platform. Then, the *read_html* function was used on every application ID to load their store's pages which contain tags. Among other things, these tags allow us to see whether a game is a VR application. The IDs of the VR application were saved into a vector.

Then, the actual scraping started. According to the documentation of Steam, the *GET* function was used in R [34]. Due to the large amount of data, only English and the most recent 1,000 reviews were scraped per game. Naturally, if a game did not have at least 1,000 reviews, all were scraped. Overall, 1,635,919 reviews were scraped. Out of them, 79.02% were positive and 20.98% were negative.

C. Data analysis

After the scraping was finished, the data analysis process started. First, the number of words was calculated with the regular expression of '[\\w\']+'. This also required the *stringr* package, and the *str_count* function in it. The number of words was also grouped by review type.

To count the occurrences of each word, the *tm* package was used. First, a corpus was created of the textual parts found in the reviews. Then, the text was manipulated as follows: words were converted to lower cases; English stopwords, punctuation, and numbers were removed; whitespaces were stripped; and lastly, the words were stemmed using Porter's algorithm [35]. Afterward, the word occurrences could be calculated. The *ggplot2* package was used to plot the most frequent words.

When investigating the word correlations with frequent words, the *tidyverse*, *tidytext*, *widyr*, *igraph*, and *ggraph* packages were used. After empirical testing, the following parameters were used. First, the words had to occur at least once in at least 100 reviews per game. Then, by using pairwise correlation, the coefficients were assessed among all possible combinations. For plotting, in the case of positive reviews the correlation coefficient was set to at least 0.7, while in the case of negative reviews, it was set to at least 0.6. Otherwise, the figures would have become unreadable.

Letter-value plots were used to better visualize the distribution of data. These types of plots were designed for large datasets and they allow us to see more reliable estimates beyond the quartiles [36]. After them, the upper and lower eights are shown, then upper and lower sixteenths, and so on until the process reaches a stopping criterion. A 95% confidence interval is determined around each letter-value. If an overlap is found between this confidence interval and the previous letter-value, then the current letter-value and the following values are not plotted anymore. Each interval is shown with darker and smaller boxes. Due to the large number of data, a graphical method was used to check whether each dataset followed Gaussian distribution. Thus, quantile-quantile (Q-Q) plots were used. Due to the results found in the previously mentioned plots, the non-parametric Wilcoxon rank sum test was used when comparing word numbers and playtimes between review types [37]. Similarly, when assessing the correlation between playtime and review length, the Spearman rank correlation method was used [38].

III. Results

This section consists of three subsections. In the first one, the length of reviews is analyzed between review types. In the second one, the playtimes when writing reviews are examined. The correlation between playtimes before reviewing and review type is also investigated in it. Lastly, in the third subsection, the most frequent words and their associations are analyzed.

A. Word numbers between review types

The next step was to compare the reviews grouped by whether they were positive or negative. There were 1,292,758 positive reviews and 343,161 negative ones in the dataset. In both types, there were several reviews that had a length of zero. Still, it can be seen that the negative reviews contained more words than positive ones. The median, mean, and standard deviations of negative reviews were 36; 82.1; and 136.73 words, respectively. On the contrary, their respective words were 18; 53.68; and 107.83 in the case of the positive reviews. The longest negative review was 2,667 words long, while the longest positive one was 2,214 words long. Overall, as can be seen, negative reviews were longer than the positive ones. The distribution of word numbers can be observed in Fig. 1.

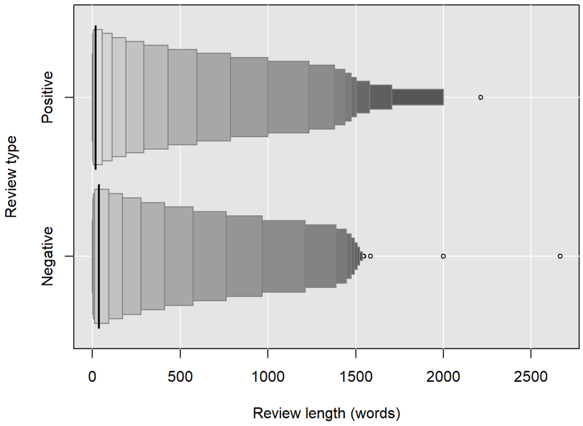


Fig. 1. The distribution of words grouped by review type.

Before using statistical tests to compare the review types to each other, it was also assessed whether word numbers follow Gaussian distribution in the dataset. Due to the large number of data, the previously mentioned Q-Q plots were used for this purpose. As can be seen in Fig. 2, the data did not follow Gaussian distribution.

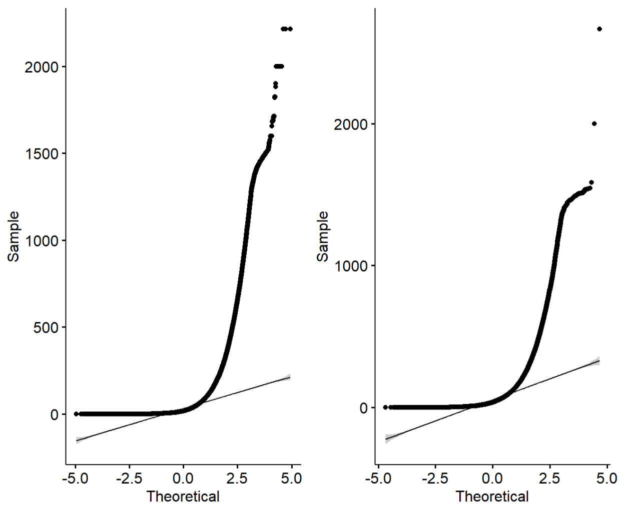


Fig. 2. Q-Q plots of word numbers in positive (left) and negative reviews (right).

Afterward, the Wilcoxon rank sum test was used to see whether the lengths of these review types were significantly different from each other. The results of the test show that the negative ones were significantly longer, $W = 1.6979 \cdot 10^{11}$; $p < 2.2 \cdot 10^{-16}$. It was also assessed with regression analysis methods whether the review type had an effect on review length. The results in Table I show that an effect exists since positive reviews were significantly shorter on average.

TABLE I
RESULTS OF THE REGRESSION ANALYSIS

	Estimate	Standard error	t value	$Pr(> \|t\|)$
Intercept	82.0994	0.1955	419.9	$< 2 \cdot 10^{-16}$
Positive review	-28.4220	0.2200	-129.2	$< 2 \cdot 10^{-16}$

As can be seen in Table I, the average length significantly differs between the two review types. On average, positive reviews contain significantly fewer words than negative ones. The mean difference between the two review types is 28.4220 words.

B. Playtime between review types

Next, the playtime of users between the two review types were investigated. According to the results, players review VR games after playing for an average of 2,172 minutes. The median value is 281 minutes, and the standard deviation is 13,502.37 minutes. Clearly, most players tend to play a VR game for several hours before leaving a review for it. It should be noted that the largest playtime before reviewing was 2,867,784 minutes. That would mean 47,796.4 hours. In the case of positive reviews, the mean, median, and standard deviation values are 2278; 335; and 13,478.66 minutes, respectively. Regarding negative reviews, their respective values are 1,773; 99; 13,584.04 minutes. As can be observed, negative reviews are written much earlier than positive ones. The distribution of playtimes between review types can be seen in Fig. 3.

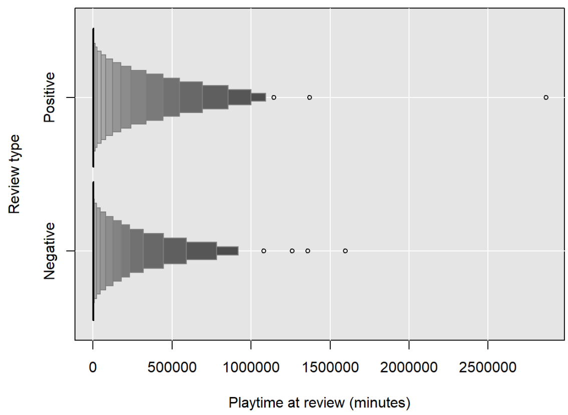


Fig. 3. The distribution of playtime before reviewing grouped by review type.

Similarly, to word numbers, Q-Q plots were used to assess whether playtimes followed Gaussian distribution. These plots can be observed in Fig. 4.

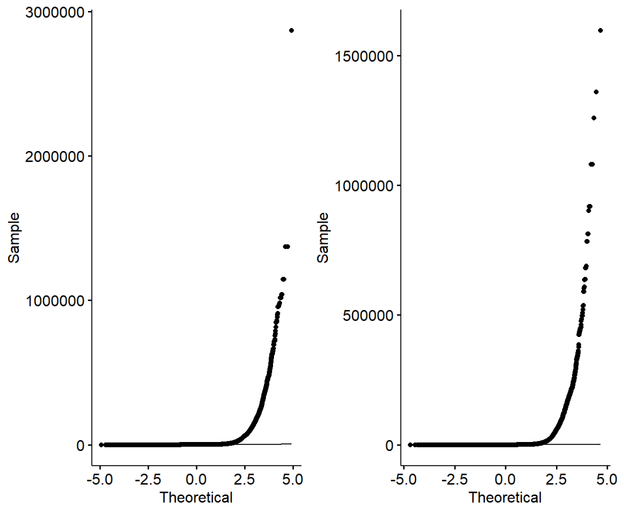


Fig. 4. Q-Q plots of playtimes in positive (left) and negative reviews (right).

As the distribution was non-Gaussian, the Wilcoxon rank sum test was used to compare the playtimes between the two review types. The results show that there is a significant difference between them, $W = 2.8911 \cdot 10^{11}; p < 2.2 \cdot 10^{-16}$. As previously, regression analysis methods were used to see whether review type had a significant effect on playtimes before reviewing. According to the results presented in Table II, a significant effect exists. Positive reviews were written significantly later than negative ones.

TABLE II
RESULTS OF THE REGRESSION ANALYSIS

	Estimate	Standard error	t value	$Pr(>\|t\|)$
Intercept	1,772.74	23.09	76.76	$< 2 \cdot 10^{-16}$
Positive review	505.04	25.98	19.44	$< 2 \cdot 10^{-16}$

It is shown in Table II that level is significance is strong in each case. On average, positive reviews are written significantly later than negative with a difference of 505.04 minutes. This would mean that those who play VR for more hours are more likely to leave a positive review.

However, it is imperative to understand whether a correlation exists between playtime and review length. Therefore, the next step was to assess this. First, this relationship was assessed on the whole dataset. However, no correlation was found, $r_s(1{,}635{,}917) = -0.046;\ p < 2.2 \cdot 10^{-16}$. When only positive reviews were analyzed in this regard, no correlation was found as well, $r_s(1{,}292{,}756) = -0.040;\ p < 2.2 \cdot 10^{-16}$. Then, negative reviews were analyzed, however, they presented a similar relationship between the two variables, $r_s(343{,}159) = 0.106; p < 2.2 \cdot 10^{-16}$. Still, it can be noted that the correlation coefficient was somewhat stronger in this case.

C. The most frequent words and their associations

Next the most frequent words were assessed in both positive and negative reviews. In case of positive ones, they are shown in Figs. 5 and 6.

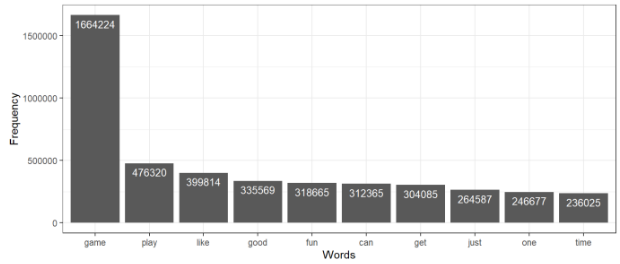


Fig. 5. The most frequent words in positive reviews.

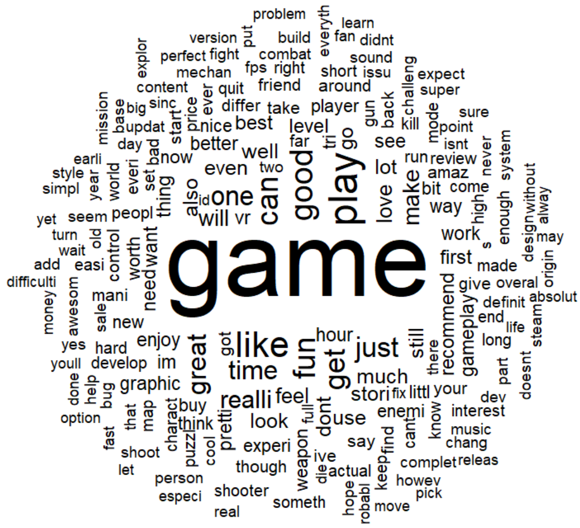


Fig. 6. Word cloud of words in positive reviews.

As can be observed in Figs. 5 and 6, the most frequent word in positive reviews is “game” with 1,664,224 occurrences. The second most frequent was “play” with 476,320 occurrences. Judging from the words, players in positive reviews mention their experiences (e.g. “fun”, “love”, “good”, “great”, “awesome”, “enjoy”), game difficulty, game mechanics, and that they recommend the game. Interestingly, even within positive reviews, there were mentions of bugs, suggesting that while these issues were acknowledged, they did not detract from the overall positive experiences of players. This indicates that the identified bugs may not have severely impacted the overall enjoyment and satisfaction derived from the games, as evidenced by the overwhelmingly positive sentiments expressed by users.

The most frequent words in negative reviews were examined next. They can be observed in Figs. 7. and 8.

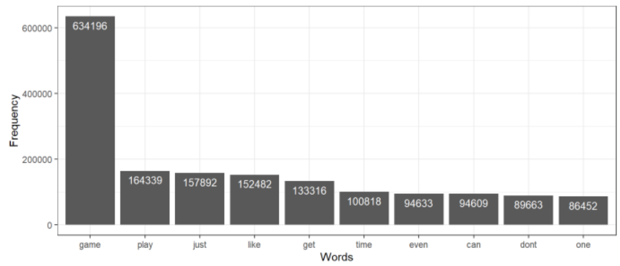


Fig. 7. The most frequent words in negative reviews.

Fig. 8. Word cloud of words in negative reviews.

Similarly to positive reviews, “game” and “play” were the two most frequent words in negative reviews, with 634,196 and 164,339 occurrences, respectively. This highlights the emphasis on gaming experience and active engagement within the VR gaming environment, albeit in a negative context. However, in negative reviews, the word “don’t” emerged as a prominent term, replacing “time” in the top 10 most frequent words, indicating dissatisfaction among users. Correspondingly, the negative sentiment was reflected in the increased frequency of words such as “poor”, “bad”, “bug”, “problem”, and “crash”, signifying a focus on technical problems and performance issues encountered by players. Furthermore, the presence of the word “server” within these reviews suggests that network-related or device-related problems contributed to the negative experiences reported by users. Additionally, the mention of “fix” underscores the imperative for developers to address the identified bugs within the games. Notably, the inclusion of the word “refund” indicates that users expressed a desire for reimbursement due to the unsatisfactory nature of their experiences. Players also showed concerns related to game mechanics, specifically pertaining to combat and weapons, as well as references to gameplay difficulty, as evidenced by the inclusion of the word “hard”. Despite many negative sentiments, a few positive words were also present, albeit less frequently. This indicates a nuanced and varied spectrum of user feedback. Moreover, the words in Fig. 8 revealed the emergence of profanity in the reviews, with the f-word being notably prevalent, showing negative emotions and frustration expressed within the negative reviews.

Naturally, these words alone do not give enough context to understand the strengths and weaknesses of VR games. Thus, as was mentioned in Section II, word associations of commonly occurring words were also examined. The results of this investigation can be observed in Fig. 9 in case of both review types.

When examining positive reviews, the following observations could be made. Among others, the words “voice acting” can be commonly found in these reviews. The cases are similar with the words “single player”, “trial error”, “learning curve”, “motion sickness”, and “paced fast”. Similarly, “extremely”, “absolutely”, “fantastic”, “favorite”, “excellent”, “totally”, “highly” and “recommended” have a strong correlation. Interestingly, profane words are still highly correlated with each other in positive reviews. Even though there were some negative word associations like between “motion” and “sickness”, they were not negative enough to warrant a negative review.

Negative reviews however, present different correlations between words. The correlated word with “money” is “waste”, indicating that players regretted buying the game. Similarly, “alt” correlates with “f4”, suggesting that players frequently quit the games due to anger or frustration. The word

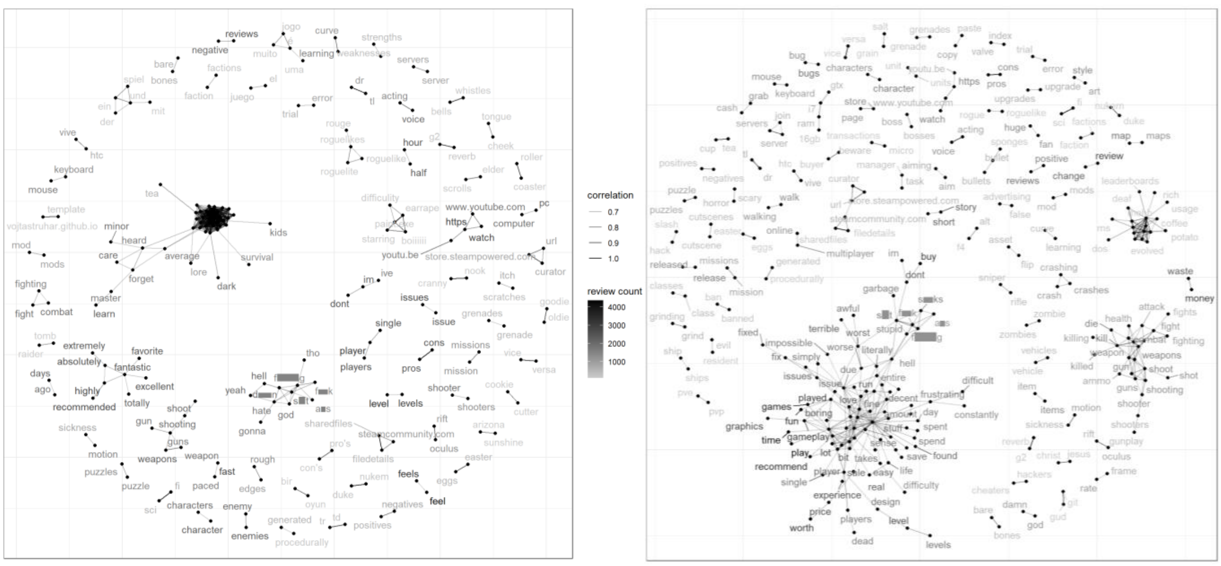


Fig. 9. Frequent words and their associations in positive (left) and negative reviews (right).

"advertising" correlates with "false", showing that even the ads for these games contained false information. "Short" is commonly found associated with "story", indicating that the length of these games was also a problem. Looking at the second largest cluster of word correlations on the right, it can observed that game mechanics regarding shooting with guns, or simply battle mechanics were badly designed, resulting in in-game character death. When examining the largest cluster of word correlations, the following can be stated. Firstly, those words appeared in most negative reviews. Secondly, the cluster of words is full of profanity and negative adjectives. Thirdly, these words can be associated with the game's difficulty or boring gameplay, bad level design, bugs/issues, and not recommending the games. Interestingly, "sale" is also mentioned, indicating that the games should be bought during a sale as they were not worth their full asking price.

IV. Discussion

The research questions were answered by the findings which provided an understanding of VR game reviews. Thus, in this section, the implications of the results are discussed.

The first research question focused on understanding the length of VR game reviews. The results indicate that negative reviews tend to be longer on average compared to positive ones. This suggests that players express their dissatisfaction in greater detail. This is similar to the results presented by Lin et al. [29]. In their study, they compared the length of reviews between various video game genres. Although they did not assess VR games, our results show that a similar phenomenon exists in their case.

The second research question revolved around the relationship between playtime and word numbers in the reviews. While no correlation was found, there was a significant difference in when players wrote the reviews. Positive reviews are written much later than negative ones. This raises interesting questions about player motivations. Negative reviews are written earlier due to frustration or disappointment, while positive ones are written after a more extensive gameplay session that strengthens positive impressions. These results are also similar to the ones concluded by Lin et al. when they examined review times between video game genres [29]. Our results show that VR games show a similar phenomenon in this regard.

The third research question focused on the understanding of word frequencies and their associations in the reviews. The results provided valuable contextual information about player experiences. Positive reviews highlight the players' appreciation for narrative elements, gameplay, and game mechanics. On the contrary, negative reviews often emphasized dissatisfaction with purchasing decisions and they showed the importance of how players evaluate the value of money. Similarly, players were critical of game length and narrative depth. The strong correlations between certain words, particularly profanity and negative adjectives, highlights the emotional and critical nature of negative reviews. These findings show that players use specific language to express their frustrations with game mechanics, and certain design elements.

This study holds significant implications for both VR game developers and researchers. For developers, gaining insights into the correlation between review length and playtime can serve as a valuable foundation for creating new strategies that are aimed at enhancing player engagement and immersion. By addressing the highlighted concerns within negative reviews, such as game mechanics, level design, and fixing the identified bugs, developers can actively contribute to an improved player satisfaction and experience. Moreover, the word associations within the reviews present an opportunity to see the preferences and priorities of players. With this understanding, developers can optimize and refine aspects of their games that resonate positively with players, while simultaneously identifying and solving issues that create negative sentiments. Furthermore, the identification of prevalent terms within positive reviews, a roadmap can be provided for developers. With it, they can reinforce the strengths of their games and bolster player satisfaction as well. Using these results, developers can more effectively tailor their future projects to align more closely with the preferences and expectations of their player base.

Naturally, this study has its limitations. Firstly, only English reviews were scraped, and the limit of reviews was 1,000 per VR game. Secondly, the analysis was restricted to textual content. Thirdly, only Steam reviews were investigated. Future research could expand the analysis by including other digital video game distribution platforms and other sources of feedback, such as forum discussions or social media conversations. Also, sentiment analysis techniques and natural language processing algorithms could be included to provide a deeper understanding of emotions in the reviews.

V. Conclusions

In this study, we analyzed VR game reviews to gain a deeper understanding of player experience. We investigated 1,635,919 textual reviews comprising both positive and negative reviews from the Steam digital video game distribution platform. Our goal was to understand review length, playtime before reviewing, and frequent words as well as their associations regarding VR games.

The results within this study show that negative reviews are written significantly earlier and contain more words than positive ones. This suggests that players tend to write about their dissatisfaction as early as possible and they express their concerns in greater depth. On the contrary, positive reviews contain concise language. These positive reviews also highlight the significance of narrative elements, gameplay, and learning experiences. Negative reviews, on the other hand, range from monetary concerns to gameplay mechanics, and even false advertising. The resulting word correlations provided a glimpse into players' experience.

In conclusion, our results can provide recommendations for developers and researchers alike. By understanding the relationships between review length, playtime, and word frequency as well as their associations, richer player experiences can be created in the realm of VR.

Acknowledgment

This work has been implemented by the TKP2021-NVA-10 project with the support provided by the Ministry of Culture and Innovation of Hungary from the National Research, Development and Innovation Fund, financed under the 2021 Thematic Excellence Programme funding scheme.

References

[1] G. Wersényi, "Musings on the Metaverse and on how the Internet of Digital Reality encompasses future developments of the Internet," *Infocommunications Journal*, vol. 15, no. Special Issue: Internet of Digital & Cognitive realities, pp. 21–28, 2023, DOI: 10.36244/ICJ.2023.SI-IODCR.4.

[2] F. Erdős, R. Németh, and B. Firuza, "Virtual Teamwork in Gamified 3D Environment," *Infocommunications Journal*, vol. 15, no. Special Issue: Internet of Digital & Cognitive realities, pp. 15–20, 2023. DOI: 10.36244/ICJ.2023.SI-IODCR.3.

[3] F. Faiqotuzzulfa and S. A. Putra, "Virtual reality's impacts on learning results in 5.0 education: A meta-analysis," *International Transactions on Education Technology (ITEE)*, vol. 1, no. 1, pp. 10–18, 2023, DOI: 10.33050/itee.v1i1.172.

[4] S. G. Izard, J. A. Juanes Méndez, and P. R. Palomera, "Virtual reality educational tool for human anatomy," *Journal of Medical Systems*, vol. 41, no. 5, 2017, DOI: 10.1007/s10916-017-0723-6.

[5] M. A. Rojas-Sánchez, P. R. Palos-Sánchez, and J. A. Folgado-Fernández, "Systematic literature review and bibliometric analysis on virtual reality and education," *Education and Information Technologies*, vol. 28, no. 1, pp. 155– 192, 2023, DOI: 10.1007/s10639-022-11167-5.

[6] P. Smutny, "Learning with virtual reality: a market analysis of educational and training applications," *Interactive Learning Environments*, pp. 1–14, 2022, DOI: 10.1080/10494820.2022.2028856.

[7] B. Xie, H. Liu, R. Alghofaili, Y. Zhang, Y. Jiang, F. D. Lobo, C. Li, H. Huang, M. Akdere, C. Mousas, and L.-F. Yu "A review on virtual reality skill training applications," *Frontiers in Virtual Reality*, vol. 2, 2021, DOI: 10.3389/frvir.2021.645153.

[8] D. W. Carruth, "Virtual reality for education and workforce training," in *2017 15th International Conference on Emerging eLearning Technologies and Applications (ICETA)*, Stary Smokovec, Slovakia, 2017, pp. 1–6, DOI: 10.1109/ICETA.2017.8102472.

[9] N. Gasteiger, S. N. van der Veer, P. Wilson, and D. Dowding, "How, for whom, and in which contexts or conditions augmented and virtual reality training works in upskilling health care workers: Realist synthesis," *JMIR Serious Games*, vol. 10, no. 1, p. e31644, 2022, DOI: 10.2196/31644.

[10] M. Matsangidou, F. Frangoudes, M. Hadjiaros, E. Schiza, K. C., Neokleous, E. Papayianni, M. Avraamides, and C. S. Pattichis "'Bring me sunshine, bring me (physical) strength': The case of dementia. Designing and implementing a virtual reality system for physical training during the COVID-19 pandemic," *International Journal of Human-Computer Studies*, vol. 165, no. 102840, p. 102840, 2022, DOI: 10.1016/j.ijhcs.2022.102840.

[11] S. Z. A. Ansari, V . K. Shukla, K. Saxena, and B. Filomeno, "Implementing virtual reality in entertainment industry," in *Cyber Intelligence and Information Retrieval*, Singapore: Springer Singapore, 2022, pp. 561–570, DOI: 10.1007/978-981-16-4284-5_49.

[12] F. Rega and D. Saxena, "Free-roam virtual reality: A new avenue for gaming," in *Advances in Augmented Reality and Virtual Reality*, Singapore: Springer Singapore, 2022, pp. 29–34, DOI: 10.1007/978-981-16-7220-0_3.

[13] G. Burdea and P. Coiffet, *Virtual Reality Technology*. Nashville, TN: John Wiley & Sons, 1994.

[14] I. Horváth, Á. B. Csapó, B. Berki, A. Sudár, and P. Baranyi, "Definition, background and research perspectives behind 'Cognitive Aspects of Virtual Reality' (cVR)," *Infocommunications Journal*, vol. 15, Special Issue: Internet of Digital & Cognitive realities, pp. 9–14, 2023, DOI: 10.36244/ICJ.2023.SI-IODCR.2.

[15] C. V ogel and A. Esposito, "Interaction Analysis and Cognitive Infocommunications," *Infocommunications Journal*, vol. 12, no. 1, pp. 2–9, 2020, DOI: 10.36244/ICJ.2020.1.1.

[16] H. Mclellan, "Cognitive issues in virtual reality," *Journal of Visual literacy*, vol. 18, no. 2, pp. 175–199, 1998, DOI: 10.1080/23796529.1998.11674538.

[17] S. Hedegaard and J. G. Simonsen, "Extracting usability and user experience information from online user reviews," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013, DOI: 10.1145/2470654.2481286.

[18] H.-N. Kang, "A study of analyzing on online game reviews using a data mining approach: STEAM community data," *International Journal of Innovation, Management and Technology*, pp. 90–94, 2017, DOI: 10.18178/ijimt.2017.8.2.709

[19] C. Chambers, W.-C. Feng, S. Sahu, and D. Saha, "Measurement-based characterization of a collection of on-line games," in *Proceedings of the 5th ACM SIGCOMM conference on Internet measurement – IMC '05*, 2005.

[20] I. Busurkina, V. Karpenko, E. Tulubenskaya, and D. Bulygin, "Game experience evaluation. A study of game reviews on the steam platform," in *Communications in Computer and Information Science*, Cham: Springer International Publishing, 2020, pp. 117–127, DOI: 10.1007/978-3-030-65218-0_9.

[21] X. Li, Z. Zhang, and K. Stefanidis, "A data-driven approach for video game playability analysis based on players' reviews," *Information (Basel)*, vol. 12, no. 3, p. 129, 2021, DOI: 10.3390/info12030129.

[22] S. Ahn, J. Kang, and S. Park, "What makes the difference between popular games and unpopular games? analysis of online game reviews from steam platform using word2vec and bass model," *ICIC Express Letters*, vol. 11, pp. 1729–1737, 2017.

[23] K. Wase, P. Ramteke, R. Bandewar, N. Badole, B. Kumar, and M. Bogawar, "Sentiment Analysis of product review," *International Journal of Innovations in Engineering and Science*, vol. 3, pp. 8–13, 2018.

[24] G. Andreev, D. Saxena, and J. K. Verma, "Impact of review sentiment and magnitude on customers' recommendations for video games," in *2021 International Conference on Computational Performance Evaluation (ComPE)*, Shillong, India, 2021, pp. 992–995, DOI: 10.1109/ComPE53109.2021.9752380.

[25] V. de Souza Goncalves and A. A. Macedo, "Similarities and divergences in electronic game review texts," in *2020 19th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)*, Recife, Brazil, 2020, pp. 145–153, DOI: 10.1109/SBGames51465.2020.00028.

[26] E. Guzman and W. Maalej, "How do users like this feature? A fine grained sentiment analysis of app reviews," in *2014 IEEE 22nd International Requirements Engineering Conference (RE)*, Karlskrona, Sweden, 2014, pp. 153–162, DOI: 10.1109/RE.2014.6912257.

[27] C. Lu, X. Li, T. Nummenmaa, Z. Zhang, and J. Peltonen, "Patches and Player Community Perceptions: Analysis of No Man's Sky Steam Reviews," in *DiGRA'20-Proceedings of the 2020 DiGRA International Conference*, 2020, pp. 1–24.

[28] Y. Yu, B.-H. Nguyen, F. Yu, and V.-N. Huynh, "Esports game updates and player perception: Data analysis of PUBG steam reviews," in *2021 13th International Conference on Knowledge and Systems Engineering (KSE)*, Bangkok, Thailand, 2021, pp. 1–6, DOI: 10.1109/KSE53942.2021.9648670.

[29] D. Lin, C.-P. Bezemer, Y. Zou, and A. E. Hassan, "An empirical study of game reviews on the Steam platform," *Empirical Software Engineering*, vol. 24, no. 1, pp. 170–207, 2019. DOI: 10.1007/s10664-018-9627-4.

[30] B. Dean, "Steam usage and catalog stats for 2023," *Backlinko*, 17-Mar-2023. [Online]. Available: https://backlinko.com/steam-users. [Accessed: 19-Aug-2023].

[31] E. Blancaflor and J. M. G. San Miguel, "Analyzing Digital Game Distribution in Gaming Industry: A Case Study," in *Proceedings of the 2nd Indian International Conference on Industrial Engineering and Operations Management*, 2022, pp. 674–681.

[32] Valve Corporation, "Documentation home page (steamworks documentation)," *Steamgames.com*. [Online]. Available: https://partner.steamgames.com/doc/home. [Accessed: 19-Aug-2023].
[33] Valve Corporation, "User reviews – get list (steamworks documentation)," *Steamgames.com*. [Online]. Available: https://partner.steamgames.com/doc/store/getreviews. [Accessed: 19-Aug- 2023].
[34] Valve Corportation, "ISteamApps interface (steamworks documentation)," *Steamgames.com*. [Online]. Available: https://partner.steamgames.com/doc/webapi/ISteamApps. [Accessed: 19-Aug-2023].
[35] C. J. Van Rijsbergen, S. E. Robertson, and M. F. Porter, "New models in probabilistic information retrieval," in *British Library. (British Library Research and Development Report*, London, 1980.
[36] H. Hofmann, H. Wickham, and K. Kafadar, "Letter-value plots: Boxplots for large data," *Journal of Computational and Graphical Statistics*, vol. 26, no. 3, pp. 469–477, 2017,
DOI: 10.1080/10618600.2017.1305277.
[37] H. B. Mann and D. R. Whitney, "On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other," *The annals of mathematical statistics*, vol. 18, no. 1, pp. 50–60, 1947.
[38] C. Spearman, "The proof and measurement of association between two things," *American Journal of Psychology*, vol. 15, no. 1, p. 72, 1904,
DOI: 10.2307/1412159.

Tibor Guzsvinecz is an Assistant Professor at the Faculty of Information Technology of the University of Pannonia, Hungary. He obtained his PhD degree in the field of Information Science in 2021. His research areas are data science, gamification, human- computer interaction, spatial ability, and virtual reality. He authored 69 scientific publications which include 19 journal papers, and 41 studies found in various conference proceedings. At the moment, he is an editor of a book project about gamification. He is an IEEE member since 2023.

Judit Szűcs is an Assistant Professor at the Faculty of Information Technology of the University of Pannonia, Hungary. She received her PhD degree in 2021 in Computer Science at the University of Szeged. Her main interests include binary tomography, image processing, and computer vision. She has 30 scientific publications which include 6 journal papers, and 19 studies found in various conference proceedings. She is an IEEE member since 2023.

Special Issue
of the Infocommunication Journal

Design and Evaluation of Abstract Aggregated Avatars in VR Workspaces

György Persa*, and Ádám B. Csapó†‡

Abstract—**Avatars are commonly used in digital platforms to provide a visual representation of individual users to each other. Generally, avatar design in the past has focused on achieving visual fidelity and realism of social interactions. In this paper, we broaden the concept of avatars to incorporate displays using an abstract visual language and conveying information on aggregated, interpersonal information from the perspective of the digital platform as a whole. We propose a general design methodology for such aggregated avatars, and also introduce and experimentally evaluate an aggregated avatar which we have developed on the MaxWhere VR platform. Results are promising in that users were able to discern several key states of the avatar and correctly associate them with the correct virtual reality scenarios in a statistically meaningful way.**

Index Terms—**virtual reality, avatar design, abstract avatars, aggregated avatars**

I. Introduction

The word '*avatar*' originates from Sanskrit, and refers to the meaning of 'embodiment', or a 'divine being made of flesh' [1]. In the rapidly evolving fields of virtual reality and metaverse technology, it evokes a very specific connotation of a human-like visual representation that can convey real-time information on the appearance, activities and even the mental state of a given user [2], [3].

Despite this clear picture of what an avatar is, or should be, it is nevertheless worth noting that avatars can take many shapes and forms, depending on the specific information they are used to represent, the visual and cognitive fidelity of their representation, and even depending on the context in which they are used. At one 'corner' of this multi-dimensional spectrum, an avatar can appear as a simple pictogram, showing a static photo of the user and indicating their presence on the social / virtual platform. At another 'corner' of the spectrum, highly visually realistic avatars are used that digitally re-represent the user's facial and head movements based on real-time tracking. In a third 'corner' of the spectrum, one can consider digital environments that involve many users interacting at once, instead of focusing on one-on-one, or small-group social interactions. In the case of a personalized virtual reality working environment, helping users to either streamline their digital workflows, or to focus on personal productivity and personal growth in an engaging environment would be the most relevant feature. In such scenarios, avatars that provide motivation to users by reflecting back to them key insights on their activities as a kind of meta-cognitive feature of the platform might be most valuable. Such meta-cognitive functionality for an avatar can also in fact be considered in cases where a platform has many users and the aggregate statistics reflecting the way in which the platform is being used will be of more interest than the interaction patterns of any given user. In this scenario, each user might contribute in a small way to the overall behavior of the avatar.

In this paper, we broaden the traditional understanding of avatars to include representations that mirror not just the specifics of a single user to others but also provide a holistic view of the digital environment to the user. This includes the state of the digital platform and the interactions of the user and possibly numerous other users with the platform. By easing the rigid definitions of what an avatar represents and to whom, we propose the possibility of previously unexplored avatar types, which we label as 'abstract avatars' and 'aggregated avatars'. We further posit that these types of avatar hold significant potential in our evolving digital landscape, where human and digital interactions are deepening not only in the short run but also in a more sustained, co-evolutionary manner as often described in the literature on cognitive infocommunications and cognitive aspects of virtual reality [4], [5]. To validate this concept, we describe a possible design methodology for such avatars, develop an example implementation, and perform validation of this implementation to demonstrate the viability of this approach.

The paper is structured as follows. In Section II, we consider different dimensions along which avatars can be qualified, and define novel categories of avatars based on these dimensions, including abstract and aggregated avatars. In Section III, we describe the design steps we have taken to design an aggregated avatar based on emotional features. In Section IV, we describe a framework within which we validate the design of our aggregated avatar through test subjects. Finally, Sections V and VI present details on the experimental design we have used, and the results we have obtained.

II. Conceptualizing novel types of avatars

A. Visual and cognitive fidelity

As discussed in the previous section, the visual representations used in avatars can range from simple, static pictograms to highly accurate visual representations of specific users [2], [6]–[10]. Based on this, it is possible to consider the *visual*

* Doctoral School of Multidisciplinary Engineering Sciences, Széchenyi István University, Győr Hungary
† Corvinus Institute for Advanced Studies & Institute of Data Analytics and Information Systems, Corvinus University of Budapest, Hungary
‡ Hungarian Research Network, Budapest, Hungary
E-mail: persa.gyorgy@sze.hu
E-mail: adambalazs.csapo@uni-corvinus.hu

DOI: 10.36244/ICJ.2024.5.11

fidelity of an avatar, in terms of how it relates to the user it represents in a visual sense.

At the same time, avatars can also be assessed in terms of their behaviors, whether in terms of how naturally they behave in social interactions [11]–[15], or in terms of how well they reflect the key aspects of user interactions, thereby contributing to increased productivity. Based on this, we have defined the term *cognitive fidelity* as follows [16]:

Definition 1: The **cognitive fidelity** of an avatar is a an assessment, both qualitative and potentially quantitative, of how well an avatar reflects the state of an underlying process, viewed from the attendant benefits to the cognitive capabilities of the users to whom it appears.

We note that the term cognitive fidelity has been used in other contexts, for example, to describe how well the use of a virtual tool corresponds to users' actions and possible choices [17], [18]; or in other cases, to describe the ecological validity of an environment from the standpoint of cognitive tasks being carried out inside it [19]. In the above definition, we focus instead on the mental capabilities of the user to whom the avatar is displayed.

As an example, let us consider a flight simulator. In such an application, the designer would have a choice as to whether to represent the co-pilot of the user as a photo-realistic human whose body posture and head movements are fully in accordance with normal social interactions, or instead, to use more abstract representations that can offer useful feedback as to what is going wrong or what aspects of flight control require attention. In some cases, even the feedback that "something is going wrong" can be highly relevant and can improve cognitive performance.

B. Abstract and Aggregated Avatars

When the main focus is on cognitive rather than visual fidelity, questions of information modeling and representation mapping naturally arise.

In terms of information modeling, it is often the case that parameters that are not so directly linked with physical reality (for example, statistics on different interaction types, or other parameters closely related to the application scenario) need to be communicated through the avatar.

In terms of finding a useful representation, the goal is to map the information to be represented onto visual (and perhaps other sensory) channels in a way that is recognizable and intuitively meaningful to users. From the theory of cognitive infocommunication (CogInfoCom) channels, if we consider avatars to be analogous to CogInfoCom messages, then the information mapping types of direct (both low-level and high-level), as well as indirect (including structural, co-stimulation or scenario-based) can be especially relevant [4], [20].

Based on these considerations, we have introduced the following definitions of abstract and aggregated avatars, with the intention of encompassing a broader scope of potential avatar designs [16]. Here, we propose a slightly revised version of this definition so as to focus – in the case of abstract avatars – solely on the language of the avatar, irrespective of the application scenario:

Definition 2: An **abstract avatar** is an avatar representation that relies on an abstract, low-level visual language involving the use of dynamic shapes and colors without reference to anthropomorphic or zoomorphic concepts.

Definition 3: An **aggregated avatar** is an abstract avatar that is used to display the features of an impersonal set of interactions and contextual events in a computational environment.

On the one hand, aggregated avatars are capable of providing users with a mirror of their position within a cooperative process, as seen through the perspective of the digital environment, instead of merely presenting data about other users. Conversely, due to its largely impersonal character, aggregated avatars can also serve to inform users about the comprehensive 'status' of a platform like a workplace or a virtual reality setting. This includes collective, environmental concepts such as the 'vibrancy of the surroundings' or the 'enthusiasm level of participants'. These aspects are somewhat influenced by the conduct of the users involved, yet they cannot easily be broken down into the distinct individual behaviors of each user.

C. Design principles for the development of abstract and aggregated avatars

Based on the dichotomy of information and representation modeling, the design of an abstract or aggregated avatar can be broken into the following steps:

1) Answering the question of what information types to model, including potential range of numerical values. In case the state to be modeled includes (fuzzy) categorical features as well, e.g. fuzzy modeling can be used to convert states into numerical fuzzy membership values.
2) Answering the question of how the resulting variables can be mapped onto an intermediate language suitable for driving discernible and interpretable avatar behaviors. One example of such an intermediate language can be a valence-arousal based emotional model, which is capable of representing distinct emotional states that can be recognized with relative ease by many users.
3) Finally, defining the visual features of the abstract avatar that can be used to drive its behavior, and mapping onto them the states of the intermediate language described in the previous point.

III. Design of an Emotionally Based Aggregated Avatar

We have given design principles for aggregated avatars in previous sections, aiming to enhance cognitive fidelity by finding suitable mappings between cognitive attributes of users and visual features of avatars. In our prior works, we created and implemented an abstract avatar inspired by ethology that can express emotional states [21]. We used this avatar design to present the aggregated space data to users.

A. Emotion displaying agent

The abstract agent that we designed to display emotions has two distinct parts: A colored sphere that can change

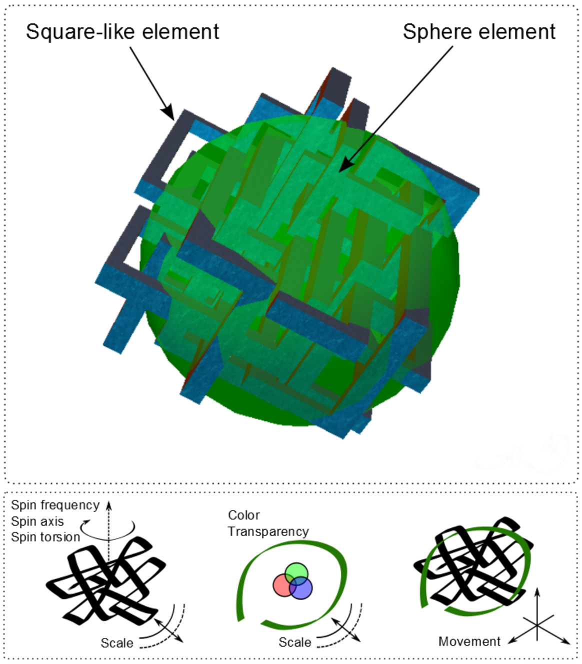


Fig. 1. The appearance and degrees of freedom of the emotional display agent

its transparency and size; and a cube-like structure that has different colors on each side and resembles a maze 1. Degrees of freedom included the position of the component, size of elements separately, color and transparency of the sphere element and rotation speed and angle of the square-like element. This setup provided us with the possibility of evoking associations with emotional states [22].

A possible way to illustrate the aggregated state of a virtual environment is to use the emotion displaying agent as a collective representation of the space. The agent acts as an abstract aggregated avatar that reflects the overall mood of the virtual space through its visual features. This approach requires a mapping function that can translate the aggregated state of the space into the visual parameters of the agent's appearance and behavior.

B. Mapping emotions to agent

Based on the evaluation of expressive features of the Emotion Display Agent, we designed a continuous mapping which can construct the look of the Emotion Display Agent for any combination of emotions. For this purpose we used the well-known Valence-Arousal model of emotions. This psychological framework describes how emotions are represented in two dimensions: valence refers to the positive or negative quality of an emotion, while arousal refers to its intensity or activation level.

The main goal with the mapping was to produce the kinds of appearances in the case of different valence-arousal value compbinations that have already been associated with corresponding emotion (happiness, sadness, fear, etc.) by users in the previously cited study [22]. Thus, we designed the following rules to achieve this:

- High valence and high arousal drove the avatar into a state where it expressed happiness by growing in size, with the sphere element exceeding the inner maze shape and appearing in a yellowish color while rotating at a relaxed pace
- High arousal with low valence caused the avatar to display anger: the sphere element turned red and shrunk while the maze element exceeded the sphere in size and rotated with high speed and non-linear easing.
- Low valence with low arousal caused the avatar to display a sad expression. It rotated evenly, moved away from the camera and shrunk in every dimension while the maze elements greatly exceeded the size of the sphere. The color looked approximately pale purple in this state
- High valence with low arousal reflected a bored or sleepy state of the avatar. To express this emotion, the maze element rotated at a low speed but unevenly, while the sphere grew big and appeared in a purple-like color
- In its natural state the avatar looked relaxed, with a green colored sphere slightly exceeding the size of the maze element. Rotation and easing were adjusted to normal pace in this case.

The top-right corner of Figure 2 shows several examples of how arousal and valence were mapped onto the avatar parameters.

C. Mapping aggregated interactions to emotions

As described in [5], modern infocommunication platforms are evolving to encompass a wide variety of novel interfaces, including virtual reality interfaces, AI-driven interfaces and distributed Web 3.0 applications. In this context, new kinds of aggregated information types are emerging which, when communicated to users in a way that represents the system as a whole, can provide intuitive feedback on parameters such as how active, how overloaded, or how quiet the system is – whatever the case may be.

In the context of virtual reality, parameters such as number of users, activity level of users in terms of moving around in the space, exploring localized subsets of the space, or interacting with others in the space could be of interest.

The specific mapping functions that are used will naturally depend on the aggregated data and the goal of the system in communicating it to users. In Section IV-B, we present a detailed example of one specific application we have developed to test the aggregated avatar concept.

IV. A Test Framework for Evaluating Abstract and Aggregated Avatars

To evaluate the effectiveness of the abstract aggregated avatar, we developed a test framework in the MaxWhere VR application. The main objective was to obtain quantitative data

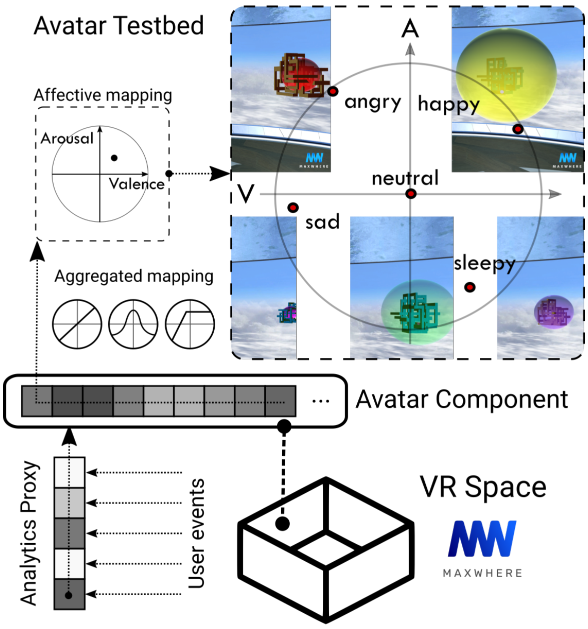


Fig. 2. Architectural diagram of testbed framework used for investigating aggregated avatar in MaxWhere VR.

on how well the participants could understand and differentiate the aggregated state of the virtual space. The basic idea was to design a test in which we could present various user interaction scenarios in the virtual environment and also display the state of the abstract aggregated avatar. We could then test whether a participant could identify the relationship between a user interaction and the corresponding changes in the visual appearance of the avatar by either swapping or keeping aligned the interaction videos and the avatar recordings, and repeating this idea with several different kinds of interactions. The overall structure of the framework is shown in Figure 2.

A. MaxWhere VR

MaxWhere VR is an innovative virtual reality platform that enables users to create and experience immersive virtual environments for various applications. MaxWhere VR offers a versatile and customizable framework that can suit different needs and goals. Users can access a rich library of 3D scenes or import their own content to design their own virtual worlds with custom functionalities. It can be used for presenting, showcasing, collaborating or research purposes as well.

Unlike other platforms, MaxWhere VR offers a unique feature called the "Where Object Model" (WOM), which is similar to a Document Object Model (DOM) used in web programming, but for 3D spaces. The WOM allows users to access and manipulate the properties and behaviors of 3D objects using Javascript code. Users can also create and load components, which are reusable pieces of code that can enhance the functionality and interactivity of the 3D spaces. With MaxWhere, users can design and program their own 3D spaces in a flexible and dynamic way, using powerful programming interfaces.

The platform also provides users with so-called "smartboards", which are essentially 2D display panels that can be used to show web content in a customizable size, position and orientation within the 3D space.

Due to its ability to customize the environment MaxWhere VR is convenient for designing scientific experiments in virtual spaces. Adding custom functionality to virtual spaces allows researchers to perform an experiment directly in the virtual space or collect data from the application for further analysis. The business logic of the experiment can be implemented as a component and can be easily added to any MaxWhere space. Many examples of such environments can be found in the literature (see e.g. [20], [23]–[25]).

B. Analytics Proxy and Testbed Component

MaxWhere provides several proxies to aid component developers accessing higher level information about 3D objects or application state. Proxies are modules which use the MaxWhere engine WOM API to implement functionalities and encapsulate them into a dedicated interface. Interfaces defined this way can be made available for MaxWhere component developers by exposing functions to the WOM API thus making them available for any MaxWhere user. Using this structural possibility we developed a Space Analytics MaxWhere Proxy with the aim of reporting about the user interaction in the space conveniently. The proxy provides basic information about the virtual environment, such as the sum axis aligned bounding box (AABB) of the space calculated from each visual object located in the space. Furthermore event listeners can be registered via the interface which reports about the following events:

- User (camera) movement in the space
- User changes between 3D and menu of the space
- User discovers a 3D object by orbiting around it
- User interact with a smartboard in the space with mouse cursor

The reporting methods for these events can be adjusted on the proxy. Users can specify the reporting interval if the continuous event sending is not desirable. Also, the behavior for unchanged states can be modified as the user can choose whether the idle system should report the identical events or not.

Using the generic Space Analytics Proxy we implemented a custom Testbed MaxWhere component for creating statistics from the space data. For the calculations we used a customizable time window (default is 40 seconds) in which each type of data is aggregated. We construct the following properties from the gathered information:

- Ratio of discovered area of the space relative to the total AABB. Covered area is calculated by creating an AABB from camera positions
- User mobility: Ratio of time the user spent moving
- Ratio of time the user spent interacting on a smartboard. Calculated from mouse cursor moves while 2D content is displayed
- Ratio of time the user spent in a smartboard, on the menu or in 3D

- Ratio of time the user spent orbiting around a 3D object

The produced statistics are then converted into the intermediate interpretation of Arousal and Valence values. For this calculation we defined the processing functions for each parameter of the aggregated Space Analytics data. We used three different transition functions: One for a high effect with a quick slope, one which increases the effect of a parameter as it nears the average value and one for a small contribution with stretched out slope. Using these transfer functions we tailored the Arousal and Valence values so that:

- High valence was associated with the statistical parameter values all being around average, which means that the user interactions are not one-sided and the space capabilities are all used for some degree.
- High Arousal values are directly but not linearly correlated with user movements and orbiting around objects, or with the discovered area increasing. Switching between 3D space and 2D menu / interaction on smartboards only had a limited effect of this value.

The calculated affective values were then used to drive the look of the aggregated avatar inside the MaxWhere space. The avatar – as summarized in section III-A - is able to express emotions via changes in its visual parameters. To map affective values derived from spatial events to the adjustable attributes of the abstract avatar we use the mappings presented in section III-B. The data flow behind the framework is presented on Figure 2.

Note that we have also implemented utilities for recording and replaying user interactions and avatar states. When recording, the avatar is temporary hidden and the user can interact with the space normally. Each interaction received on the Analytics Proxy is recorded and written into a file. The proxy is configured to report every state of the space periodically, thus a continuous sequence of states is stored as a result. Using the replay function, it is then possible to read these files at a later time and to drive the aggregated avatar with the recorded space state values. During replay the space is hidden and only the avatar is visible. Hiding the avatar or the space in these utilities was important for creating sample sequences and videos for further analysis.

V. Experimental design

In the following section we describe the details of the experiment we performed to validate the feasibility of our aggregated abstract avatar.

A. *Test videos*

We used the record features of the MaxWhere testbed component to create several pairs of test videos to evaluate the performance of the avatar. Each pair of videos included a user interaction video and an avatar behavior video. The user interaction videos captured only the scene and the interactions without the avatar, while the reactions of the avatar to these interactions were recorded from the saved log file using the replay function. During replay, we recorded the corresponding video showing only the expressions of the avatar. To facilitate comparison, we selected three different aspects of user behavior as dimensions for choosing typical interaction patterns for the videos. The three dimensions were as follows.

- *Discovery*: Spatial behavior when the user moves around the virtual space and discover multiple 3D objects. It consists of the space analytics data of camera movements and orbiting.
- *Manipulation*: Describes the scenario when the user interacts with the 3D objects using the Editor capabilities of MaxWhere. It includes a little bit of orbiting from the space analytics data and mostly specified by menu transitions and cursor movements. During this scenario the user opens and closes the menu multiple times, inputs values on the user interface or moving around 3D object or resize them in the space with dragging gesture of the cursor.
- *2D Operation*: Characterized by smartboard interactions. Using the space analytics data of entering or leaving a smartboard and cursor moments while a smartboard is selected for usage. This is the case of browsing the web or working on a content in smartboards.

We recorded 8 pairs of videos in total along these dimensions. For each interaction types we made two videos where only that behavior was shown (six videos in total). We also made one video where the user showed no behavior and one video where the user showed all the behaviors. The original videos were about one minute long, but we sped them up to 30 seconds for the experiment.

A frame from a video of a user interaction and the corresponding avatar behavior can be seen on Figure 3.

B. *Survey*

We conducted an experiment to test how well participants could match user interaction videos with avatar behavior videos. We have chosen the aggregated state of the virtual space to be the *”usefulness”* of user interactions. This means that the ”mood” of the affective avatar was driven by the user interaction types and based on the quality and the quantity of interactions the avatar displayed a combination of ”happy”, ”frustrated” or ”angry” state.

We created 8 video pairs in total, each consisting of a user interaction video and a corresponding avatar behavior video as described in the previous section. We then randomly selected 2 video pairs for each task and presented them to the participants in a mixed order: first a user interaction video, then another user interaction video, then an avatar behavior video, and then another avatar behavior video. The participants had to decide which of the two avatar behavior videos matched the first user interaction video they saw. In other words, they had to choose between the combinations of (1-3; 2-4) and (1-4; 2-3). This task is presented to the participants four times with different video pairs each time.

The complete experiment consisted of the following steps:

- Language selection (we prepared an English and a Hungarian version)
- First we introduced the details of the research and explained the tasks ahead

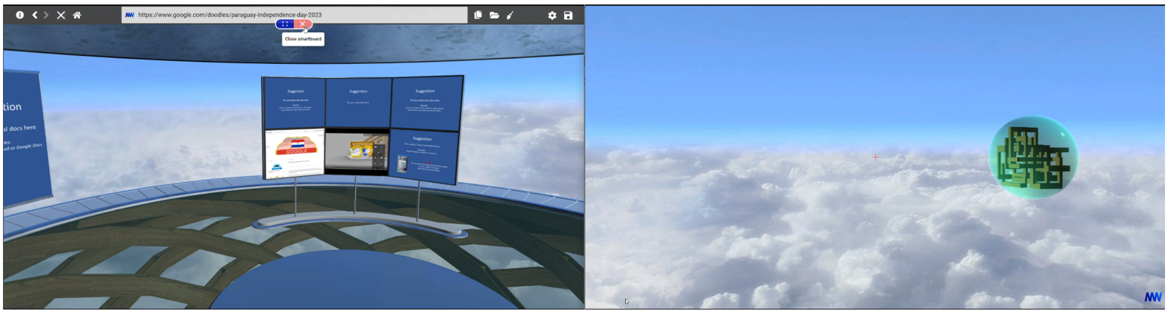


Fig. 3. Example of user interaction video (on the left) with the corresponding avatar video (on the right) showing a 2D operation scenario.

- Collection of informed consent from the participants for anonymized data collection
- Presentation of the 4 matching tasks

The briefing text was the following:

> "*This study explores the expressive capabilities of an abstract-shaped avatar.*
> *The avatar aims to convey a summary (aggregated state) of events occurring in the virtual space. Since the avatar has an abstract shape, it can use only abstract methods (based on emotional associative techniques) to express its state. This test evaluates how well the avatar's expression is understood.*
> *In the following sections, we will present you with 4 tasks, each consisting of 4 videos. The first 2 videos in each task will show user interactions in the virtual space, while the third and fourth video will show the avatar's response to user behaviors from one of the user videos. Your goal will be to pair the first two (user) videos with the second two (avatar) videos.*
> *We used MaxWhere VR application to create the virtual environment for this experiment. This application allows users to access or upload media contents through interactive boards in the 3D space, also known as "smartboards". The user interaction patterns in this system can be classified into three types:*
> - navigation: moving in the space
> - 2D interaction: viewing content on smartboards
> - manipulation: moving, resizing smartboards
>
> *We recorded user interaction videos in this study that include these patterns. These patterns affect the "mood" of the abstract avatar, which changes the displayed overall state of the environment to combination of these values: "bored", "excited", "happy" or "frustrated". In the following tasks, the overall state will be the* "usefulness" *of user interactions, which means how much the user takes advantage of every function available in the space. For instance,*
> - a 'pointless wandering' pattern will trigger a frustrated mood,
> - a task pattern will trigger a combination of bored mood and excitement based on how repetitive is the task,
> - and a balanced use of all interaction types will trigger a happy mood.
>
> *The recordings are about 30 seconds long and each are accelerated to the same degree.*"

We also present a brief video to the participants after the introduction, which demonstrates the typical user interactions in MaxWhere VR that are described in the briefing text. This is essential for ensuring that they can recognize and differentiate the actions on user interaction videos with as much confidence as a more experienced MaxWhere user.

C. Automated data collection

The survey used an automatic recording system to store each response in a spreadsheet that was linked to the survey. The recording also triggered several custom routines that analyzed the answer and produced some statistical metrics. To do this, we created a custom scripting interface that connected the survey with the data collector modules that were integrated into the survey. The custom routines generated the survey for the next participant. This way, the next version of the test had different order and combinations of answers and videos.

The survey for the experiment is created in Google Forms linked to a Google spreadsheet to store the results. Custom routines were implemented in Google Apps Script framework to process the data. The script runs automatically after each Form submission, triggered by the framework settings.

D. Custom routines

Custom routines performed upon survey completion includes randomization of the next test and evaluation of the current one.

Video randomization is performed using the following steps in the script.

- Read out the URLs of video pairs from a spreadsheet page. Each video pair (user interaction and avatar video) is associated with an ID.
- Associate groups to the video pairs according to which user interaction type they contain
- For the first task, select the Idle video and pair it with one of the one-dimensional videos (Discovery, Manipulation, Operation2D).

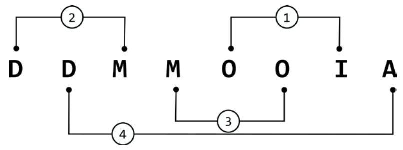


Fig. 4. Example run of the randomization routine. Numbers depicts the order of operation. Letters are the user interaction types (D - discovery, M - manipulation, O - 2D operation, I - idle, A - all)

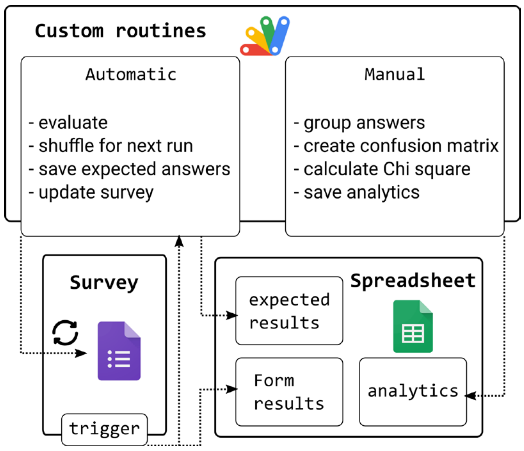


Fig. 5. Data flow of survey, connected spreadsheet and custom routines. Automatic routines run when a survey is completed. It evaluates current answers and shuffle pairs for a new run. Manual routines run on demand and create analytics.

- For second and third task, pair the least used one-dimensional videos randomly. Least used videos are determined by counting how many times a video is already used for pairing in previous tasks. (thus, it is actually forced for the second task but not for the third)
- Pair the All dimensional video with the remaining video for the last (fourth) task.
- Randomize order within pairings to alter display order of videos in the survey.

The outcome of random pairing for the tasks of the next survey is then registered in a spreadsheet. It represents the expected answers for the next submission. An example run of the randomization routine is demonstrated on Figure 4.

The form submission trigger runs the script after the connected spreadsheet has been updated with the latest responses. This allows the script to automatically check the accuracy of the matching tasks. The script retrieves the correct answers from the sheet based on the latest response.

The overall functionality of the evaluation and randomization script can be seen on Figure 5.

VI. Results

In this section we summarize the statistical analysis of the survey responses. The main objective of the analysis is to determine how well the users can match the avatar behaviors with the corresponding user interactions.

The survey has been distributed only for a smaller group at first for fine-tuning. Based on the preliminary findings described below, we modified the survey and recruited a larger group of participants through social media and technical forums. We analyzed the final responses using the custom manual methods explained in the previous section.

A. Preliminary results

For the test video pairs, we initially had a slightly different setup. In the first version of the experiment we combined 2 different interaction types as well instead of using only one, zero or all dimensions for creating test videos. The randomization algorithm also did less instructed shuffling, which made the video match tasks more challenging.

We recorded answers from 12 people (5 male, 7 female) and found that none of the interaction types produced acceptable recognisability. We also observed that the randomized task generator created complex combinations that increased the difficulty level for even the most skilled MaxWhere users.

B. Revised experiments

Based on the preliminary results we simplified the experiment tasks by reducing the number of user interaction types depicted in the test videos. This made the videos more distinguishable as the participants did not see tasks with combinations of the same user interaction type.

We also modified the shuffle algorithm in order to fix one half of the random pairings in half of the tasks. The first task has always presented the zero dimensional case and the last task had all the user interaction types for one of the video pairs. The revised video pairing is described in steps in section V-D. The new routine solved the distribution problems as well, because each video pair is used once for each participant.

The explanatory text at the beginning of the survey has also been modified in order to provide a better understanding of the upcoming tasks. We added information about MaxWhere VR in general and introduced the typical user interaction types performed in the user videos of the tasks in a bullet point list and on a short video. We emphasised that the aggregated state of the space is the ”usefulness” of user interactions in this experiment and listed several examples for mappings between the user interaction and displayed emotions of the avatar. For example, we described that a happy avatar means that the user interaction was useful, while a frustrated or bored avatar means that it was not. These sentences helped the participants to understand the situation presented on user interaction videos and emotional states shown on avatar videos. We considered this to be reasonable for our experiment, realizing that it would have been difficult to expect users to associate interactions with the avatar if they had not been attuned to the kinds of interactions that exist in the first place, or the general purpose of the avatar.

C. Statistical analysis

We recorded responses from 31 participants (8 female, 22 male, 1 undefined). Most of them were middle-aged with average age between 30 and 41. Based on the media channels we distributed the survey, we could ascertain that participants

TESTS	Discover	Manipulate	Operate2D	Idle	All	MATCHES	Discover	Manipulate	Operate2D	Idle	All
Discover	0	20	23	8	11	Discover	0	14	12	7	7
Manipulate	20	0	19	14	9	Manipulate	14	0	12	11	7
Operate2D	23	19	0	9	11	Operate2D	12	12	0	7	7
Idle	8	14	9	0	0	Idle	7	11	7	0	0
All	11	9	11	0	0	All	7	7	7	0	0
EXPECTED	Discover	Manipulate	Operate2D	Idle	All	OBSERVED	Discover	Manipulate	Operate2D	Idle	All
Discover	31	10	11.5	4	5.5	Discover	40	6	11	1	4
Manipulate	10	31	9.5	7	4.5	Manipulate	6	44	7	3	2
Operate2D	11.5	9.5	31	4.5	5.5	Operate2D	11	7	38	2	4
Idle	4	7	4.5	15.5	1.00E-10	Idle	1	3	2	25	0
All	5.5	4.5	5.5	1.00E-10	15.5	All	4	2	4	0	21

Fig. 6. Tables used for extracting input data for Chi Square test. Tests table consists of participation of interaction types in tests. Matches table shows the successful matches of the user interaction types. Expected and Observed tables are calculated for Chi Square test using the basic assumption that user interaction types can be recognized only as good as 50%

Confusion / Occurences	Discover	Manipulate	Operate2D	Idle	All
Discover		0.7	0.522	0.875	0.636
Manipulate	20		0.632	0.786	0.778
Operate2D	23	19		0.778	0.636
Idle	8	14	9		0
All	11	9	11	0	

Fig. 7. Confusion and occurrence matrix of the matching results. The matrix has two parts: the lower left part shows the total number of appearances of each user interaction type combination, and the upper right part shows the rate of successful matches for each combination. Darkness of background color of cells indicates the significance of values

CHIDIST	Discover	Manipulate	Operate2D	Idle	All
p values	0.141611638	0.02256998518	0.3981787884	0.0193337095	0.3849561179

Fig. 8. P values of Chi Square test. The result shows significant difference for Idle and Manipulate state and unconfirmed difference for Discover and Operate2D.

had adequate computer skills and at least half of them were familiar with 3D video games.

We performed various statistical tests and analyses to assess how well our avatar can convey the aggregated state of the space during different types of user interaction. To obtain meaningful statistics, we first grouped the data by dimensions of user interaction type. We counted the successful matches from the tasks for every participant and registered a successful recognition of the aggregated state for each successful match. Since the matching tasks involved two different dimensions each time, we recorded success or failure for both of the involved interaction types.

We constructed two basic tables from the extracted data: The total number of participations and number of successful matches for each user interaction dimension. Both tables were derived from the spreadsheet that contained the pairing information about each test run. Using the pairing information alone we could construct the participation matrix. The automated evaluation script registered the success for each task of each test run, which was used to construct the correct matches matrix. Results of extracted tables can be seen on Figure 6.

Based on the participation and correct match tables we constructed a confusion matrix. This matrix shows how well the participants matched the items correctly in each task. It helps us to evaluate how well the expressive ability of our avatar performs for different types of user interactions. We obtained the accuracy by dividing the number of correct matches by the number of total appearances. Figure 7 shows the confusion matrix we generated. The Idle and Discover scenario had the highest type combination score. This makes sense because the avatar shows anger in one case and boredom in the other.

For our statistical test we used Chi square test to see how the aggregated avatar performed. A Chi square test is a statistical method that can be used to test the association between pairs of categorical variables. The test compares the observed frequencies of each category with the expected frequencies under the null hypothesis of no association. The p-value of the test is the probability of obtaining a test statistic as extreme or more extreme than the observed one, under the null hypothesis. A small p-value indicates that there is strong evidence to reject the null hypothesis and conclude that there is an association between the variables – in our case, that users peformed better than randomly in the case of at least one category.

As a baseline, or null hypothesis we used the assumption that the avatar does not help in recognition of the aggregated state of the virtual environment. This means that we suggested that the participants choose randomly during the matching tasks giving us 50% success rate for each user interaction types.

Therefore, we set the expected table for Chi square test reflecting this statement. Each value in the matrix has been calculated by taking the half of the total occurrences of the given type combination. We used the matrix diagonal to store the summed result for a given dimension. For example, the total expected value for Discover user interaction type is the sum of each test combination containing Discovery divided by two, which is 62 / 2. Note, that we also simplified the case of zero values in the expected matrix to keep the calculation straightforward. We replaced zeros (which would cause a divide error during the upcoming calculations) with 1E-10 values. The expected table can be seen in lower left side of Figure 6.

We obtained the observed values for the Chi square test from the correct match matrix. In the diagonal we use the number of observed successful matches for a given interaction type which is calculated by the sum of correct matches for a given dimension in each combination it participates. For non-diagonal cells we used the number of *incorrect* choices of a given interaction type combination as the expected values describes the choice of the given value *despite of* the correct answer. We calculated these values by subtracting the correct match value for each combination from the total number of occurrences. For example, see the first two cell of the Observed matrix on Figure 6. The observed total of successful matches of Discover dimension is 14 + 12 + 7 + 7 = 40. The total number of times when participants mistakenly chose Manipulate instead of Discover when these two dimensions were combined is 20 - 14 = 6

Using the composed matrices we could calculate the P values of the Chi square test for each user interaction type as shown on Figure 8. We used 5% for the statistical threshold to reject the null hypothesis.

Based on the results we can state that Idle and Manipulate state show extreme deviation from the values picked by chance (Manipulate 2.2%, Idle 1.9%), thus the effect of our avatar for choosing the correct matching were statistically significant. With other words it is extremely improbable that the results are unrelated to the avatar setup. All and Operate types showed no real deviation from the randomly selected answers thus they could not produce evidence for the usefulness of the avatar. Discover type however shows high probability for existing effect of the avatar without being considered as significant result (14% chance for no connection).

VII. Conclusions

In this paper, we broadened the scope of the avatar concept to include multi-user avatars using an abstract visual language, referred to as aggregated avatars. We argued that such avatars could be useful in multi-user platforms such as VR-based working environments. We proposed a design methodology and developed a reference implementation of the aggregated avatar concept on the MaxWhere VR platform.

Based on our statistical analyses we can conclude that two major types of user interactions (idle state, and rearrangement of spatial content) could be very well recognized by users with the use of our abstract aggregated avatar. A third type of interaction, referring to users moving around in the space, were also likely to be recognized by users at a higher success rate than chance. Other dimensions provided a promising but statistically not relevant result – although in the case of the "All" state, this could have been due to the fact that this state included some examples of all other interactions. Based on these results, we conclude that the proposed avatar design could be applicable to some contexts, and is worthy of further study and refinement.

VIII. Acknowledgement

Project no. C1015653 has been implemented with the support provided by the Ministry of Culture and Innovation of Hungary from the National Research, Development and Innovation Fund, financed under the KDP-2020 funding scheme.

This research is contributing to project no. 2021-1.1.4-GYORSÍTÓSÁV-2022-00081 that has been implemented with the support provided by the Ministry of Culture and Innovation of Hungary from the National Research, Development and Innovation Fund, financed under the 2021-1.1.4-GYORSÍTÓSÁV funding scheme.

The research presented in this paper was also supported by the Hungarian Research Network (HUN-REN), and was partly carried out within the HUN-REN Cognitive Mapping of Decision Support Systems research group.

References

[1] L. de Wildt, T. H. Apperley, J. Clemens, R. Fordyce, and S. Mukherjee, "(Re-) Orienting the Video Game Avatar," *Games and Culture*, vol. 15, no. 8, pp. 962–981, 2020. DOI: 10.1177/1555412019858889.

[2] M. E. Latoschik, D. Roth, D. Gall, J. Achenbach, T. Waltemate, and M. Botsch, "The Effect of Avatar Realism in Immersive Social Virtual Realities," in *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, pp. 1–10, 2017. DOI: 10.1145/3139131.3139156.

[3] F. Sibilla and T. Mancini, "I Am (Not) My Avatar: A Review of the User-Avatar Relationships in Massively Multiplayer Online Worlds," *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, vol. 12, no. 3, 2018. DOI: 10.5817/CP2018-3-4.

[4] P. Baranyi, A. Csapo, and G. Sallai, *Cognitive Infocommunications (CogInfoCom)*. Springer, 2015. DOI: 10.1007/978-3-319-19608-4.

[5] I. Horváth, Á. B. Csapó, B. Berki, A. Sudár, and P. Baranyi, "Definition, Background and Research Perspectives Behind 'Cognitive Aspects of Virtual Reality'(cVR)," *Infocommunications Journal*, no. SP, pp. 9–14, 2023. DOI: 10.36244/ICJ.2023.SI-IODCR.2.

[6] D. Aneja, D. McDuff, and S. Shah, "A High-Fidelity Open Embodied Avatar with Lip Syncing and Expression Capabilities," in *2019 International Conference on Multimodal Interaction*, pp. 69–73, 2019. DOI: 10.1145/3340555.3353744.

[7] D. Roth, J.-L. Lugrin, D. Galakhov, A. Hofmann, G. Bente, M. E. Latoschik, and A. Fuhrmann, "Avatar Realism and Social Interaction Quality in Virtual Reality," in *2016 IEEE Virtual Reality (VR)*, pp. 277– 278, IEEE, 2016. DOI: 10.1109/VR.2016.7504761.

[8] T. Alldieck, M. Magnor, W. Xu, C. Theobalt, and G. Pons-Moll, "Detailed Human Avatars from Monocular Video," in *2018 International Conference on 3D Vision (3DV)*, pp. 98–109, IEEE, 2018. DOI: 10.1109/3DV.2018.00022.

[9] L. Hu, S. Saito, L. Wei, K. Nagano, J. Seo, J. Fursund, I. Sadeghi, C. Sun, Y.-C. Chen, and H. Li, "Avatar Digitization from a Single Image for Real-Time Rendering," *ACM Transactions on Graphics (ToG)*, vol. 36, no. 6, pp. 1–14, 2017. DOI: 10.1145/3130800.31310887.

[10] J. L. Ponton, V. Ceballos Inza, L. Acosta, A. Rios, E. Monclús, and N. Pelechano, "Fitted Avatars: Automatic Skeleton Adjustment for Self- Avatars in Virtual Reality," *Virtual Reality*, vol. 27, pp. 1–20, 07 2023. DOI: 10.1007/s10055-023-00821-z.

[11] C. Pedica and H. Högni Vilhjálmsson, "Spontaneous Avatar Behavior for Human Territoriality," *Applied Artificial Intelligence*, vol. 24, no. 6, pp. 575–593, 2010.DOI: 10.1080/08839514.2010.492165.

[12] A. Kleinsmith, P. R. De Silva, and N. Bianchi-Berthouze, "Cross-Cultural Differences in Recognizing Affect from Body Posture," *Interacting with computers*, vol. 18, no. 6, pp. 1371–1389, 2006. DOI: 10.1016/j.intcom.2006.04.003.

[13] M. Fabri, D. J. Moore, and D. J. Hobbs, "The Emotional Avatar: Non-Verbal Communication between Inhabitants of Collaborative Virtual Environments," in *Gesture-Based Communication in Human-Computer Interaction: International GestureWorkshop, GW'99 Gif-sur-Yvette, France, March 17-19, 1999 Proceedings*, pp. 269–273, Springer, 1999. DOI: 10.1007/3-540-46616-9_24.

[14] J. Yang, R. T. Marler, H. Kim, J. Arora, and K. Abdel-Malek, "Multi-Objective Optimization for Upper Body Posture Prediction," in *10th AIAA/ISSMO multidisciplinary analysis and optimization conference*, p. 4506, 2004. DOI: 10.2514/6.2004-4506.

[15] R. Rivu, D. Roth, F. Alt, and Y. Abdelrahman, "The Influence of Avatar Personalization on Emotions in VR," *Multimodal Technologies and Interaction*, vol. 7, p. 38, 03 2023. DOI: 10.3390/mti7040038.

[16] G. Persa and Á. B. Csapó, "A Fuzzy Driven Framework for the Cognitive Modeling and Quantitative Analysis of Aggregated Avatars," in *CINTI-MACRo 2022*, pp. 000 143–000 150, IEEE, 2022. DOI: 10.1109/CINTI-MACRo57952.2022.10029574.

[17] P. S. Moyer, G. Salkind, and J. J. Bolyard, "Virtual Manipulatives used by K-8 Teachers for Mathematics Instruction: The Influence of Mathematical, Cognitive, and Pedagogical Fidelity," *Contemporary Issues in Technology and Teacher Education*, vol. 8, no. 3, pp. 202–218, 2008.

[18] B. Bos, "Virtual Math Objects with Pedagogical, Mathematical, and Cognitive Fidelity," *Computers in Human Behavior*, vol. 25, no. 2, pp. 521–528, 2009. DOI: 10.1016/j.chb.2008.11.002.

[19] D. Lansel, T. Kergoat, S. Bardin, and L. Deruy, "Using Virtual Reality Domain to Set Up Sports Simulation as Part of Rehabilitation," 2021.

[20] Á. B. Csapó, I. Horvath, P. Galambos, and P. Baranyi, "VRasa Medium of Communication: From Memory Palaces to Comprehensive Memory Management," in *2018 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, pp. 000 389–000 394, IEEE, 2018. DOI: 10.1109/CogInfoCom.2018.8639896.

[21] G. Persa, A. Csapo, and P. Baranyi, "CogInfoCom Systems from an Interaction Perspective – A Pilot Application for EtoCom," *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol. 16, no. 2, pp. 297–304, 2012. DOI: 10.20965/jaciii.2012.p0297.

[22] B. Korcsok, V. Konok, G. Persa, T. Faragó, M. Niitsuma, Á. Miklósi, P. Korondi, P. Baranyi, and M. Gácsi, "Biologically Inspired Emotional Expressions for Artificial Agents," *Frontiers in psychology*, vol. 9, p. 1191, 2018. DOI: 10.3389/fpsyg.2018.01191.

[23] A. Sudár and Á. B. Csapó, "Descriptive Markers for the Cognitive Profiling of Desktop 3D Spaces," *Electronics*, vol. 12, no. 2, p. 448, 2023. DOI: 10.3390/electronics12020448.

[24] B. Berki, "Does Effective Use of MaxWhere VR Relate to the Individual Spatial Memory and Mental Rotation Skills?," *Acta Polytechnica Hungarica*, vol. 16, no. 6, pp. 41–53, 2019. DOI: 10.12700/APH.16.6.2019.6.4.

[25] G. Stankov, B. Nagy, *et al.*, "Eye Tracking Based Usability Evaluation of the MaxWhere Virtual Space in a Search Task," in *2019 10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, pp. 469–474, IEEE, 2019. DOI: 10.1109/CogInfoCom47531.2019.9089922.

György Persa This author obtained his MSc degree in 2009 from Pázmány Péter Catholic University in Budapest.
Following his graduation, he conducted scientific research at the Institute for Computer Science and Control of the Hungarian Academy of Sciences.
In 2020, he began his doctoral studies at Széchenyi István University in Győr. His primary research focus is on design and implementation in the areas of virtual reality, 3D applications, avatars, human-machine interfaces and cognitive computing.

Ádám B. Csapó obtained his PhD degree at the Budapest University of Technology and Economics in 2014. From 2016 to 2022, he has worked as an associate professor at the Széchenyi István University in Győr; and, since 2022, he continued his academic journey as as an associate professor at Óbuda University, Budapest, Hungary. Since September 2023, he has held the position of associate professor at Corvinus University of Budapest. Dr. Csapó's research focuses on cognitive infocommunication channels in virtual collaboration environments, i.e. enabling users to communicate with each other and with their spatial surroundings in novel and effective ways. At the same time, he has been involved in the development of assistive technology for the visually impaired, as well as in the development of a commercial VR platform. Dr. Csapó has over 50 publications, including 1 co-authored book and 20 journal papers, and has actively participated in the organization of numerous international conferences and special issues.

Cognitive Aspect of Emotion Estimation of a Driver

Jinshan Luo, Yuki Okaniwa, Natsuno Yamazaki, Yuko Hiramatsu, Madoka Hasegawa, *Member, IEEE* and Atsushi Ito *Member, IEEE*

Abstract—**Despite rapid advancements in the automotive industry, traffic safety risks persist. Addressing this challenge requires innovative driver assistance technologies. Common accidents result from driver inattention, fatigue, and stress, leading to issues like falling asleep at the wheel and improper acceleration and braking. Our study aims to contribute to advanced driver assistance systems that adapt to drivers' emotional needs, ultimately enhancing road safety. In this paper, we mention the result of our research to estimate drivers' emotions using sensors. For that purpose, we developed a sensor network containing sensors such as EEG, eye tracker, and driving simulator. We explored the relationship. As a result, we confirmed the relation between the driver's emotions, especially sleep conditions, driving speed, duration, and brain wave behavior.**

Index Terms—**driving support system, emotion estimation, EEG, eye tracking, sensor network, persistent homology**

I. Introduction

Amidst the rapid devolution of the automotive industry, the persistence of traffic safety risks remains a critical concern. Consequently, the development of intelligent and innovative driver assistance technologies becomes imperative. Among the factors contributing to traffic accidents, the most prevalent include falling asleep at the wheel, improper acceleration, and braking errors. These mishaps stem from driver inattention, fatigue, drowsiness, and stress, all significantly contributing to road accidents. In particular, the cumulative effects of fatigue resulting from extended periods of driving can substantially impair a driver's cognitive function and judgment, ultimately leading to drowsiness. Hence, it becomes essential to promptly detect driver fatigue and drowsiness during driving and institute appropriate measures. To address this challenge, there is a pressing need to develop support systems that encourage drivers to take breaks and enhance the in-vehicle environment to optimize comfort.

Given this backdrop, our research endeavors to objectively assess driver fatigue by gathering an extensive array of data, encompassing ElectroEncephaloGraphy (EEG), heart rate, eye movements, and driving activities, all obtained through driving simulators. We aim to scrutinize this data meticulously, with the expectation of delivering valuable insights that can enhance driver safety and mitigate the risk of traffic accidents.

It is worth noting, however, that while EEG holds promise as an excellent indicator of emotional states, practical limitations arise when attempting to employ EEG sensors during driving. Additionally, EEG sensors may provide limited accuracy in measuring neural activity related to emotions occurring beyond the upper layers of the brain. [1].

The remaining part of this paper is structured as follows. Section 2 introduces a variety of related research. We explain the research goal, problems, and objectives in Section 3. The sensors and network we used in the research are explained in Section 4. Section 5 explains the detailed experiment result, and discuss and analyze the result in Section 6. Finally, we present our conclusions in Section 7.

II. Positioning and the Related Works of the Paper

Numerous studies have been undertaken to predict driver fatigue by examining the correlation between drivers' biological signals and their eye movements.

A. Relationship between this study and Cognitive Infocommunications

This paper proposes and uses a sensor system to estimate a driver's mind state and subsequently explore the relationship between emotional states and driving behavior. This study combines artificial and natural cognitive capabilities. The whole system's new hybrid cognitive capabilities fall into the concept of Cognitive Infocommunications [2,3].

One of the branches of Cognitive Infocommunications focuses on Cognitive Mobility, which investigates the entangled combination of research areas such as mobility, transportation, vehicle engineering, social sciences, artificial intelligence, and Cognitive Infocommunications [4,5].

Thus the overall new capability of the combination of the censoring system and the driver leads to a new capability of the whole system that is to improve the driving effectiveness to avoid accidents and further car design outcomes.

This work was supported in part by JSPS Kakenhi under Grants, JP17H02249, 18K111849, 20H01278, 20H05702, 22K12598, 23H03649.

Jinshan Luo is with Toyota Customizing and Development Co., Ltd, Ashigara-gun, Kanagawa, Japan (e-mail: dc227209@s.utsunomiya-u.ac.jp).

Yuki Okaniwa is with the Faculty of Economics, Chuo University, Hachioji, Tokyo, Japan (e-mail: a21.hcsk@g.chuo-u.ac.jp).

Natsuki Yamazaki is with the Faculty of Economics, Chuo University, Hachioji, Tokyo, Japan (e-mail: a22.x3yf@g.chuo-u.ac.jp).

Yuko Hiramatsu is with the Faculty of Economics, Chuo University, Hachioji, Tokyo, Japan (e-mail: susana_y@tamacc.chuo-u.ac.jp).

Madoka Hasegawa is with the School of Engineering Utsunomiya University, Tochigi, Japan (e-mail: madoka@is.utsunomiya-u.ac.jp).

Atsushi Ito is with the Faculty of Economics, Chuo University, Hachioji, Tokyo, Japan (e-mail:atc.00s@g.chuo-u.ac.jp).

DOI: 10.36244/ICJ.2024.5.12

B. Studies Using Heart Rate and Electrocardiogram [6]

This study reported that during the transition from mental fatigue to drowsiness, the number of blinks increased and, conversely, the heart rate decreased.

C. Studies utilizing gaze angle and eye rotation angle [7]

This study used image sensors to generate computational models from gaze and eye rotation angles. As a result, it was reported that the accuracy of measuring mental workload from driving was improved from eye movements.

D. Studies Using Gaze Angle and Eye Rotation Angle [8].

This study has shown that blinking decreased with task load during driving. However, another paper [9] has found the opposite result, that blinking increases with increasing task load, and there still needs to be a clear answer regarding how it can be used.

Furthermore, numerous studies have explored the monitoring of driver emotions using sensors as a means to prevent risk-taking behavior. These investigations involved using multiple driving simulators to simulate realistic driver interactions and estimated driver emotions based on driving performance data. However, these studies did not directly detect emotional data to validate the accuracy of their measurements [10]. In [11], researchers developed a sensor network to establish a mapping relationship among various sensor data to monitor driver emotions. This approach incorporated metrics such as heart rate, skin conductance, skin temperature, and facial expressions. Nevertheless, the research did not address driving performance data closely tied to driving behavior. [12] introduced a non-intrusive emotion recognition system designed for car drivers, employing a thermal camera to enhance Advanced Driver Assistance Systems (ADAS). However, it's important to note that this system has yet to be tested in actual driving conditions, which leaves room for further exploration and validation.

Prior research efforts have explored various avenues in the realm of emotion recognition for drivers. For instance, in [13], an approach centered around facial expressions was introduced. This approach leveraged a comprehensive on-road driver facial expression dataset, encompassing diverse road scenarios and corresponding driver facial expressions during driving. Meanwhile, [14] devised a methodology that combines Local Binary Pattern (LBP) features with facial landmark features to detect driver emotions. This method further employed a supervised machine learning algorithm, specifically a support vector machine, to classify different emotions effectively.

Additionally, [15] put forward an innovative approach, introducing a custom-created Convolutional Neural Network (CNN) feature learning block to enhance the performance of an existing 11-layer CNN model. This augmentation resulted in an improved and faster R-CNN face detector capable of accurately identifying the driver's face. However, it's essential to note that these studies primarily focused on processing facial image data for driver emotion recognition. They did not delve into aspects such as body motion or explore the intricate relationship between driver emotion and driving performance data.

III. Outline of this study

A. Research Goal

This research aims to estimate drivers' emotions during driving to prevent car accidents.

B. Problems and Objectives of the research

While several studies have explored the estimation of driver emotions through means such as brain waves and other bio-signals, there are two notable challenges to consider. Firstly, relying solely on EEG may be problematic due to potential variations caused by the experimental environment, introducing an element of risk. Secondly, the practicality of measuring EEG by having drivers wear sensors while driving is a concern.

To address these issues, we aim to identify alternative sensors that are easy to use, robust, and cost-effective. We will compare these potential sensors with popular vital sensors commonly employed for health monitoring, as well as sensors integrated into vehicles but not worn by individuals. By examining these options, we can explore the feasibility of replacing EEG with more practical sensor solutions for emotion detection in a driving context.

IV. Sensors

We implemented a sensor network for driver emotion monitoring around a driving simulator. In this section, we would like to explain each sensor and outline the network.

A. Driving simulator

To achieve the goal of estimating the driver's emotions by using a drive recorder and analyzing the relation between driver emotions and behavior during driving, we collect driving performance data from the driving simulator such as speed, accelerator pedal degree, brake pedal pressure, steering angle, and distance from the start.

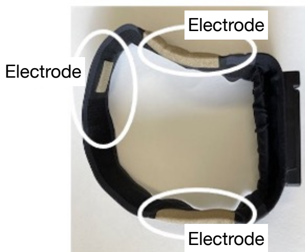


Fig. 1. EEG sensor (headband) [13]

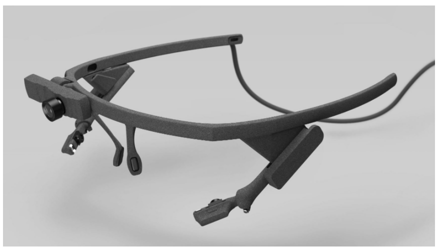


Fig. 2 Eye tracking device [16]

B. EEG sensor

We used an EEG sensor to estimate the mind state of a driver. Typically, EEG sensors are large devices with many electrodes used in hospitals, but using such devices in a car or a driving simulator is not easy [16]. So, we have developed a wearable EEG sensor that can acquire data for several hours without stress, as shown in Figure 1. This EEG sensor has BLE and sends data in real time. The sampling rate is 512 Hz. In this experiment, we used M5Stack Core2 [18] as a receiver of the EEG signal, and the data is written on an SD card in the receiver device. The EEG sensor can output the information listed in Table 1. In addition, our sensor can output two additional information: Attention (concentration, similar to Beta wave) and Mediation (relaxing, similar to Alpha wave) [19].

C. Eye tracker

We utilized an eye tracker, specifically the Pupil Core [20], to monitor eye movements accurately, as illustrated in Figure 2. This device provides us with the precise x and y coordinates of the gaze, enabling us to simultaneously capture video footage of the surrounding scenery and the movement of the eyeball.

D. Network

As shown in Figure 3, we established a sensor network for data collection, with certain components connected via Bluetooth Low Energy (BLE [21]) for real-time data transmission. In contrast, other components, such as sensors connected on Controller Area Network (CAN [22]) in the driving simulator, remained offline for security considerations.

V. Experiment and result analysis

A. Experiment Design

We set two test courses in the driving simulator. One is Tokyo Metropolitan Highway (C1), and another is a road in the center of Paris. The details of the setting are shown in Table 1. In the C1 course, we changed the brightness during driving from daytime to evening. In Paris, we used Simulation of Urban Mobility (SUMO) [23] to provide some interference to drivers, such as traffic and unexpected behavior of pedestrians.

On August 9 and 21, 2023, and November 22, 2023, two students of Chuo University (20 years old, and 2years old, owning a driver's license) drove C1 and Paris. The duration of each driving test was 45 minutes. As per our experiences, after 30 minutes, a driver starts to feel fatigued, so we set 45 minutes. In this paper, we label each trial as 20230809-C1-1, C1-2, Paris-1, Paris-2, 20230821-C1-1, C1-2, Paris-1, Paris-2, and 20231122-Paris-5.

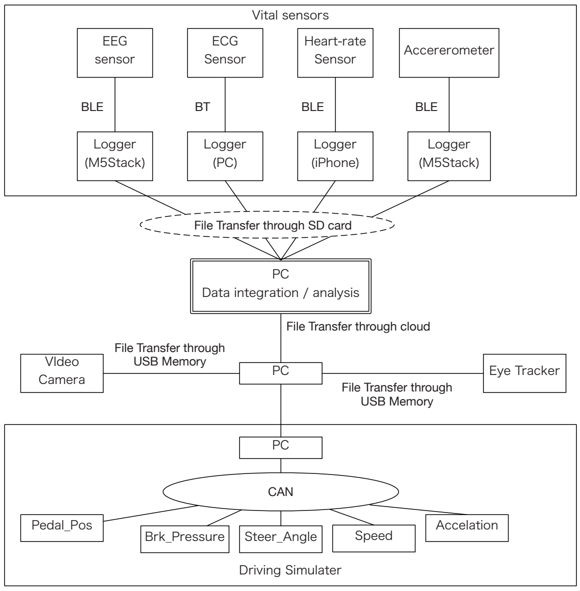


Fig. 3. The system diagram

TABLE II
Brain waves [19].

Frequency band name	Frequency	Brain states
Delta	0.5–4 Hz	Sleep
Theta	4–8 Hz	Deeply relaxed, inward focused
Alpha	8–12 Hz	Very relaxed, passive attention
Beta	12–35 Hz	Anxiety dominant, active, external attention, relaxed
Gamma	Over35 Hz	Concentration

Before and after driving, the test driver answered PANAS (The Positive and Negative Affect Schedule [24]) with 10 Positive and 10 Negative questions to record the mental state. As shown in Figure 4, the score of negative questions increased in all cases, indicating that the test driver consistently reported feeling fatigued after 45 minutes of driving.

B. Fatigue from distance and time

Figure 5 presents the relationship between the number of rounds and the duration of a single round of driving. Notably, it becomes evident that, after several rounds, the lap time increased by approximately 20%. This observation suggests

TABLE I
Test course in the driving simulator.

Course	Time/round	SUMO	Brightness change (in 45 min)
Tokyo Metropolitan Highway C1	About 10 min	No	0-10 min 4:00 PM, 10-20 min 6:00 PM 20-30 min 7:00 PM, 30-40 min 7:30 PM 40-45 min 8:00 PM (with road lighting)
Paris City Area Course	About 6 min	Traffic and Pedestrian crossing road	No

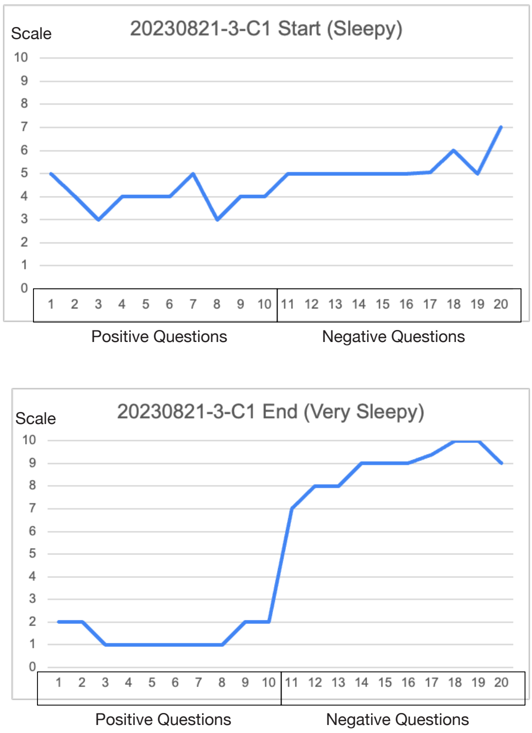


Fig. 4. Example of the result of PANAS

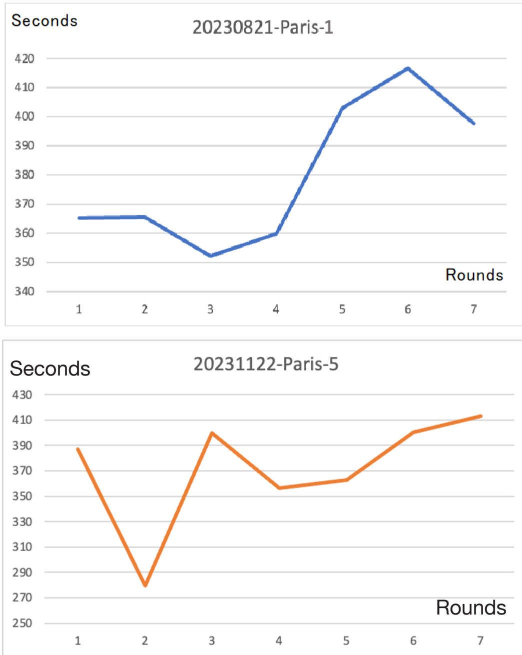


Fig. 5. Example of rap time of each round

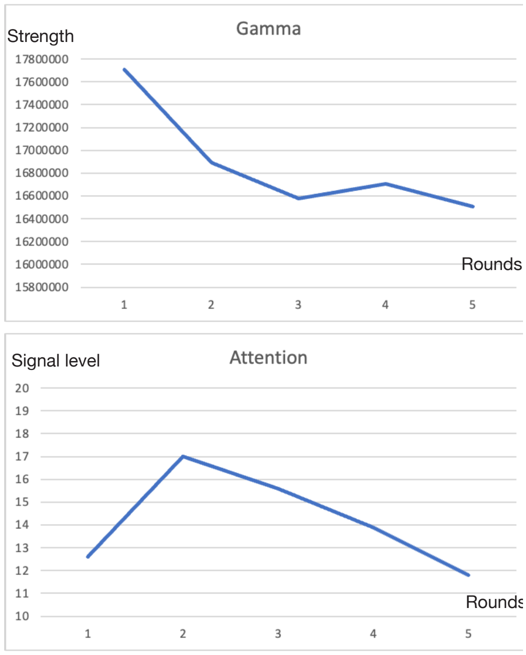


Fig.6. Trend of Gamma and Attention of 20230821-C1-1 (going down along the timeline)

that a driver's concentration tends to decline after experiencing fatigue, resulting in a slowdown in driving speed.

C. *Fatigue from changing brightness*

Figure 6 illustrates the connection between brightness levels and changes in emotion detected through EEG. As the driving rounds progress, the road becomes darker, and a corresponding decrease in driver concentration is evident. This pattern is consistent with the trends observed in C1 driving data. Conversely, during driving experiences in Paris, there was no significant decline in attention levels. Hence, we can infer that darkness has a detrimental impact on a driver's attentiveness, potentially contributing to increased fatigue and decreased concentration during nighttime driving scenarios.

D. *Relation between EEG and eye tracking*

Figure 7 illustrates the link between Gamma brainwaves and eye blinks. Our analysis of Paris data revealed a consistent pattern: when Gamma fell below 1.0E7, indicating reduced brainwave activity, the driver often lost concentration, leading to eye blinks or closures (below the red line in Figure 7). This suggests Gamma changes are a valuable indicator of tiredness, especially sleepiness, aligning with the driver's drowsiness in the latter part of the round. Additionally, this finding underscores the significance of eye tracking as a reliable method for monitoring the driver's movements and quantifying their level of fatigue while actively engaged in driving.

E. *Relation between facial recognition and eye tracking*

Facial expressions directly reflect emotions, and body motion strongly associates with emotions [25]. To estimate a driver's stress and fatigue, we harnessed the effectiveness of a widely used drive recorder, capable of capturing facial expressions and body movements. Consequently, in this study, we amalgamated driving performance data from a driving simulator with facial expression and body motion data obtained from a drive recorder for a comprehensive correlation analysis.

For facial expression recognition and head motion measurements, we utilized MediaPipe Face Mesh [26], with

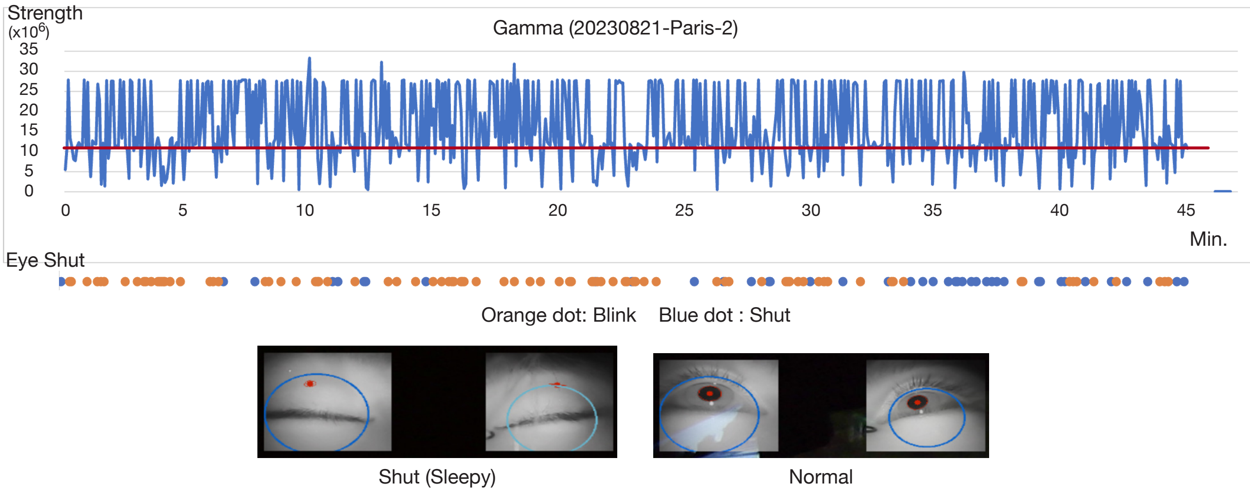


Fig. 7 Relation between Gamma wave and eye blink

prior video image preprocessing carried out using OpenCV [27]. This allowed us to visualize driving performance data concerning the accelerator pedal, brake, and steering, shedding light on their relationship with concurrent facial expressions and body motions. To assess the drive recorder's performance, we conducted a comparative analysis of emotion measurement results obtained through the drive recorder and those from an EEG sensor as our previous study [28]. This approach facilitates an evaluation of the drive recorder's effectiveness in gauging driver emotions and provides a low-cost and easily implementable method for collecting data on drivers' facial expressions through video footage was proposed.

We analyzed the connection between facial recognition and eye tracking, with a specific emphasis on the occurrence of eye closures. In Figure 8, we present six distinct patterns of facial classification observed during the experiment, which, in turn, allow us to infer four primary emotions: (1) Neutral, (2) Anxiety, (3) Boredom, and (4) Fatigue. Our primary focus lies on fatigue as it relates to the sensation of tiredness while driving. Table 3 presents the number of reported fatigue feelings and occurrences of eye closures per 5-minute intervals. The low p-value obtained from the T-Test (0.4) further validates the strong relationship between eye movements and facial expressions, supporting our hypothesis that eye tracking effectively correlates with driver emotion, particularly in instances of fatigue.

VI. Discussion

This paper aims to estimate a driver's emotion by investigating the relation between EEG, driving record, brightness, eye tracking (eye shut), and facial recognition. At this moment, we found the relation as shown in Figure 9. Unfortunately, we could not collect enough data to analyze the relation between these data and other data such as car operation (pedals, steering), heart rate, and body motion.

Firstly, we compared the facial expression estimation of fatigue by face recognition (Fig. 10) and the number of eyes

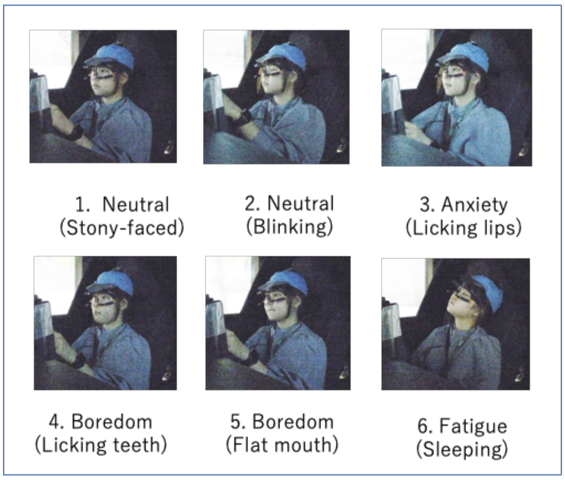


Fig.8 Six patterns from facial recognition

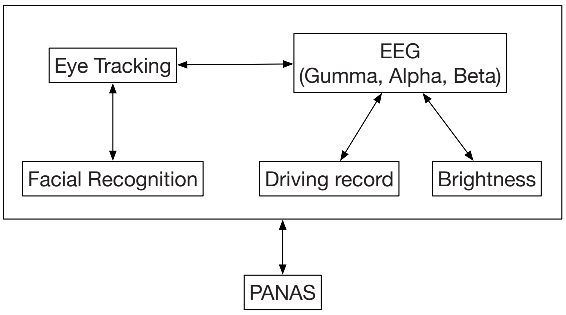


Fig. 9 Relation of information that was found from the experiment

shut (= sleepiness) by eye tracker (Fig. 11). The number of eye closures tends to increase with each round of testing. The same trend is observed on different experimental days and with different subjects. In contrast, there is a difference in the trend of facial expression estimation of fatigue even when the same

TABLE III
THE NUMBER OF FATIGUE AND EYE SHUT EVERY FIVE MINUTES.

	~05:00	~10:00	~15:00	~20:00	~25:00	~30:00	~35:00	~40:00	~45:00
Fatigue	1	5	1	1	4	5	3	10	9
Eye shut	2	2	4	0	3	5	6	13	0

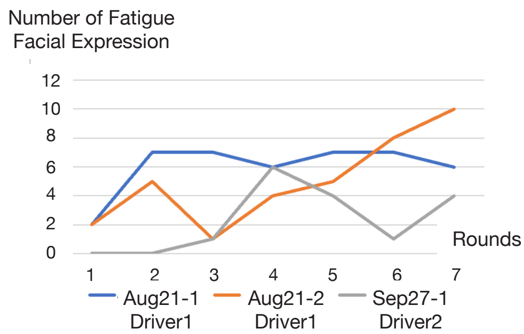


Fig.10 Number of fatigue from facial recognition (Paris, different driver)

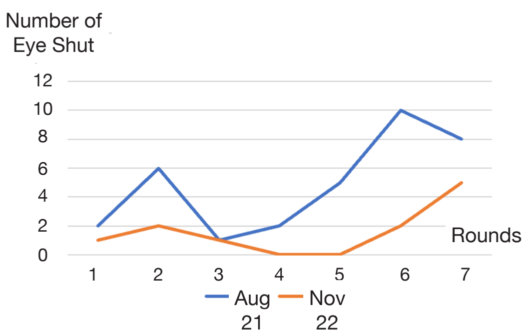


Fig.11 Number of eye shut (Paris, different driver)

subject is tested on the same day (Aug. 21-1 and Aug. 21-2 in Fig. 10). This result suggests that there is a limitation in finding fatigue estimation by face recognition.

For the lap time, as shown in Figure 5, each additional lap is getting longer. This result suggests a similar trend between eye tracking data and the number of laps (= driving time).

Then, we compared the EEG and other data. TGAM outputs raw data at 512 Hz and aggregated values of each brain wave band (Delta, Theta, Alpha, Beta, and Gamma) per second. However, it is difficult to read their interrelationships because EEG changes significantly from one second to the next. Therefore, we applied topological data analysis (TDA) [29] to analyze EEG data (Alpha, Beta, and Gamma). The basic technology of TDA is persistent homology. The following is an overview of persistent homology from [29].

"A key mathematical apparatus in TDA is *persistent homology*, which is an algebraic method for extracting robust topological information from data. To provide some intuition for the persistent homology, let us consider a typical way of constructing persistent homology from data points in a Euclidean space, assuming that the data lie on a sub- manifold. The aim is to make inference on the topology of the underlying manifold from finite data. We consider the r-balls (balls with radius r) to recover the topology of the manifold, as popularly employed in constructing an r-neighbor graph in many manifold learning algorithms. While it is expected that, with an appropriate choice of r, the r-ball model can represent the underlying topological structures of the manifold, it is also known that the result is sensitive to the choice of r. If r is too small, the union of r-balls consists simply of the disjoint r-balls. On the other hand, if r is too large, the union becomes a contractible space. *Persistent homology* [30] can consider *all* r simultaneously, and provides an algebraic ex- pression of topological properties together with their persistence over r. The persistent homology can be visualized in a compact form called a *persistence diagram* $D = \{(b_i, d_i) \in R^2 \mid i \in I, b_i \le d_i\}$, and this paper focuses on persistence diagrams, since the contributions of this paper can be fully explained in terms of persistence diagrams. Every point $(b_i, d_i) \in D$, called a *generator* of the persistent homology, represents a topological property (e.g., connected components, rings, and cavities) which appears at X_{b_i} and dis- appears at X_{d_i} in the r-ball model. Then, the *persistence* $d_i - b_i$ of the generator shows the robustness of the topo- logical property under the radius parameter. "

We used HomeCloud [31], a tool for visualizing persistence; we created a 3D graph of Alpha, Beta, and Gamma for the Paris orbit on Aug.21, 2023 (2nd trial, by driver1) and the Paris orbit on Nov.22, 2023 (5th trial, by driver2). The graph of persistence generated from the 3D data of Alpha, Beta, and Gamma is shown in Figure 12. Most of the points lie on the X=Y line, but the points away from it represent the features of the data.

In Figure 13, the envelopes are added for the big-picture view of the points. It can be seen that the shape of the envelope for each lap is similar, even though two drivers with different driving skills and different schedules are driving on different dates. This result indicates that some changes may be occurring similarly for each lap. In the future, we plan to analyze the EEG movements related to fatigue by analyzing the EEG in more detail.

VII. CONCLUSION

This study aims to estimate driver emotion by using several data that we can acquire while driving a car to prevent car accidents. For that purpose, we developed a sensor network around a driving simulator using an EEG sensor, accelerometer, heart rate sensor, and eye tracker.

The paper's novel contribution is that the result displayed a relation between the driver's emotions, especially sleepy

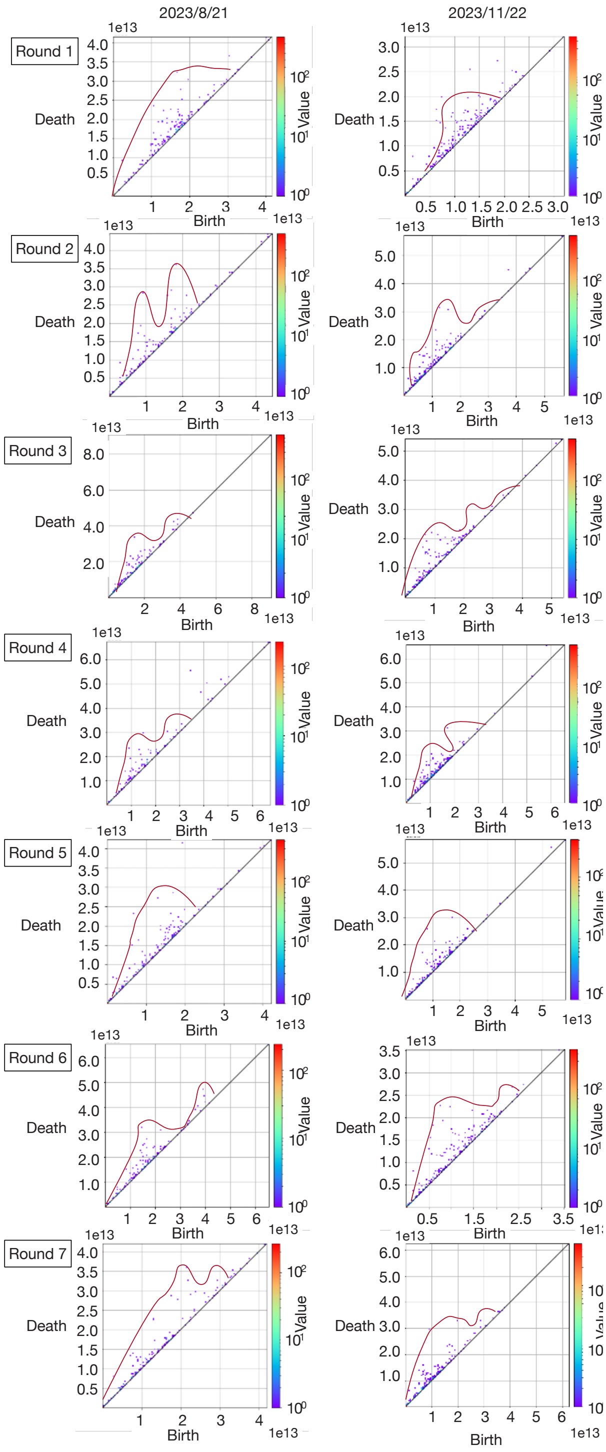


Fig. 12 Persistent homology graph of brain waves (Alpha, Beta, and Gamma)

conditions, driving speed, driving duration (=driving distance), and brain wave behavior. This result will make it possible to realize a safety-driving support technology.

However, at this moment, as motioned in section VI, our analysis was not yet sufficient because of the limited number of data. If we get enough data, we could understand the relation of information around a car to understand the driver's emotions.

However, this study showed the effectiveness of Cognitive Mobility, that is a part of Cognitive Infocommunication science.

As a further study, we need to do two things. The first is to analyze the data and information relation using new techniques such as persistent homology. The second is to acquire the Mental State Index, such as PANAS while trying to explain the mental state in language.

Our final target is to realize a system to detect the dangerous mental state related to car accidents and provide information to change the situation using a simple wearable sensor such as a smartwatch and CAN data.

Acknowledgment

The authors would like to express special thanks to Mr.Inagaki, Mr.Yuchi, Mr.Miyagawa, Mr.Sasuga, Mr.Kosuge, Ms.Tachibana, Mr.Gunji of Toyota Customizing & Development Co., Ltd. for allowing us to perform this research.

References

[1] M. Mahsa, D. Dale and K. Arshia. (2020). Electroencephalography (EEG) Technology Applications and Available Devices. *Applied Sciences*. DOI: 10.3390/app10217453.

[2] P. Baranyi and A. Csapo, “Definition and Synergies of Cognitive Infocommunications,” *Acta Polytechnica Hungarica*, vol. 9, no. 1, pp.67– 83, 2012, DOI: 10.12700/aph.

[3] P. Baranyi, A. Csapo, and G. Sallai, “Emergence of CogInfoCom in Science and Technology,” in *Cognitive Infocommunications (CogInfoCom)*, Springer, Cham, 2015, DOI: 10.1007/978-3-319-19608-4_1.

[4] Zöldy, M., & Baranyi, P. (2021). Cognitive Mobility – CogMob. In *2021 12th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)* (pp. 000 915–000 919).

[5] Zöldy, M., & Baranyi, P. (2022). The Cognitive Mobility Concept. *Infocommunications Journal*, 2023/01, pp. 35–40

[6] Gianluca Borghini, Laura Astolfi, Giovanni Vecchiato, Donatella Mattia, Fabio Babiloni, Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness, *Neuroscience & Biobehavioral Reviews*, Volume 44, 2014, Pages 58–75, ISSN 0149-7634, DOI: 10.1016/j.neubiorev.2012.10.003.

[7] Takanori CHIHARA, Fumihiro KOBAYASHI and Jiro SAKAMOTO, Estimation of mental workload during automobile driving based on eye-movement measurement with a visible light camera, *Transactions of the JSME (in Japanese)*, Vol.86, No.881, 2020

[8] Edmund Wascher, Emad Alyan, Melanie Karthaus, Stephan Getzmann, Stefan Arnau, Julian Elias Reiser, Tracking drivers’ minds: Continuous evaluation of mental load and cognitive processing in a realistic driving simulator scenario by means of the EEG, Heliyon, Volume 9, Issue 7, 2023, e17904, ISSN 2405-8440, DOI: 10.1016/j.heliyon.2023.e17904.

[9] Magliacano A, Fiorenza S, Estraneo A, Trojano L. Eye blink rate increases as a function of cognitive load during an auditory oddball paradigm. *Neurosci Lett*. 2020 Sep 25;736:135293. Epub 2020 Aug 6. PMID: 32771601. DOI: 10.1016/j.neulet.2020.135293.

[10] C. Hexuan, L. Yingzi and M. Ronald. (2007). Study on driver emotion in driver-vehicle-environment systems using multiple networked driving simulators. DSC North America – Iowa City – September North America – Iowa City – September.

[11] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[12] A. Kolli, A. Fasih, F. A. Machot and K. Kyamakya, "Non-intrusive car driver's emotion recognition using thermal camera," *Proceedings of the Joint INDS'11 & ISTET'11*, 2011, pp. 1–5.

[13] X. Huafei, L. Wenbo, Z. Guanzhong, W. Yingzhang, X. Jiyong, Z. Juncheng, L. Chengmou and G. Gang. (2022). On-Road Driver Emotion Recognition Using Facial Expression. *Applied Sciences*. 12. 807. DOI: 10.3390/app12020807.

[14] M. Patil and S. Veni, "Driver Emotion Recognition for Enhancement of Human Machine Interface in Vehicles," *2019 International Conference on Communication and Signal Processing (ICCSP)*, 2019, pp. 0420–0424.

[15] Zaman, K., Sun, Z., Shah, S.M., Shoaib, M., Pei, L. and Hussain, A., 2022. Driver Emotions Recognition Based on Improved Faster R-CNN and Neural Architectural Search Network. *Symmetry*, 14(4), p. 687.

[16] Priyanka A. Abhang, Bharti W. Gawali, Suresh C. Mehrotra, Chapter 2 – Technological Basics of EEG Recording and Operation of Apparatus, Editor(s): Priyanka A. Abhang, Bharti W. Gawali, Suresh C. Mehrotra, Introduction to EEG- and Speech-Based Emotion Recognition, Academic Press, 2016, Pages 19–50, ISBN 9780128044902, DOI: 10.1016/B978-0-12-804490-2.00002-6.

[17] Mindsall Inc., https://mindsall.com

[18] M5Stack Core2 ESP32 IoT Development Kit, https://shop.m5stack.com/products/m5stack-core2-esp32-iot-development-kit-v1-1

[19] TGAM datasheet, https://cdn.hackaday.io/files/11146476870464/TGAM%20Datasheet.pdf

[20] Pupil Core, https://pupil-labs.com/products/core/

[21] Bluetooth Technology Overview, https://www.bluetooth.com/learn-about-bluetooth/tech-overview/

[22] Controller Area Network – an overview, https://www.sciencedirect.com/topics/engineering/controller-area-network

[23] SUMO, https://eclipse.dev/sumo/

[24] David Watson;Lee Anna Clark(1999), "The PANAS-X: Manual for the Positive and Negative Affect Schedule – Expanded Form"

[25] Crane, E., Gross, M. (2007). Motion Capture and Emotion: Affect Detection in Whole Body Movement. In: Paiva, A.C.R., Prada, R., Picard, R.W. (eds) Affective Computing and Intelligent Interaction. ACII 2007. *Lecture Notes in Computer Science*, vol 4738. Springer, Berlin, Heidelberg.

[26] MediaPipe Team, Face Mesh. MediaPipe, 2022 [online].

[27] Opencv-python 4.6, 2022 [online]. https://pypi.org/project/opencv-python/.

[28] J. Luo, H. Yoshimoto, Y. Okaniwa, Y. Hiramatsu, A. Ito and M. Hasegawa, "Emotion monitoring sensor network using a drive recorder," *2023 IEEE 15th International Symposium on Autonomous Decentralized System (ISADS)*, Mexico City, Mexico, 2023, pp. 1–8, DOI: 10.1109/ISADS56919.2023.10092139.

[29] Genki Kusano, Yasuaki Hiraoka, Kenji Fukumizu, Persistence weighted Gaussian kernel for topological data analysis, *Proceedings of The 33rd International Conference on Machine Learning, PMLR* 48:2004-2013, 2016. *ArXiv*:1601.01741

[30] Edelsbrunner, H., Letscher, D., and Zomorodian, A. Topological persistence and simplification. *Discrete and Computational Geometry*, 28(4):511–533, 2002.

[31] HomCloud, https://homcloud.dev/index.en.html

Jinshan Luo received B.A. degrees in science from Chengdu University in 2016 and M.A. degrees in engineering from Utsunomiya University in 2021, respectively. He is now a Doctor's Course student at a department of engineering at Utsunomiya University since 2022. He is also a researcher for Toyota Customize and Development Co., Ltd. since 2021. His current research interests include mental state estimation, data science, image recognition, sensor network, and autonomous driving.

Yuki Okaniwa is a student of Faculty of Economics of Chuo University.

Natsuno Yamazaki is a student of Faculty of Economics of Chuo University.

Yuko Hiramatsu received the B.A. and M.A. degrees in Literature from Sophia University in 1979 and 1985, respectively. She has been on the secretariat staff of the "Study Group of School Trip" since 2006. She was an adviser for the utilization of ICT in Okinawa in the subsidy project of the Ministry of Internal Affairs and Communications in Japan from 2007 to 2009. She belongs to the Faculty of Economics at Chuo University as a lecturer since 2009 and a specially appointed professor from 2023. Her current research interests are the linguistic landscape and ICT literacy for primary education.

Madoka Hasegawa received B.S., M.S., and Ph.D. degrees in engineering from Utsunomiya University in 1994, 1996, and 1999, respectively. She is now a professor at the school of engineering, Utsunomiya University. Her current research interests include image processing, image coding, digital watermarks, and usable security. She is a senior member of IEICE (the Institute of Electronics Information and Communication Engineers, Japan) and a member of IEEE.

Atsushi Ito received B.S. and M.S. degrees from Nagoya University in 1981 and 1983, respectively. He also received a Ph.D. degree from Hiroshima City University in 2007. From 1983 to 2014, he was with the Research and Development Laboratories of KDDI Corporation, and he was a Professor at the Graduate School of Engineering of Utsunomiya University from 2014 to 2020. He has been a Professor at the Faculty of Economics of Chuo University since 2020. During 1991-1992, he was a visiting scholar at the Center for the Study of Language and Information (CSLI) at Stanford University. His current research interests include ad hoc networking, biosignal applications, and ICT-based agriculture.

Special Issue
of the Infocommunication Journal

Developing sustainable logistic strategies in the context of cognitive biases

Péter Földesi[1], and Eszter Sós[2]*

Abstract—**Cognitive biases often occur even in the decision-making process of highly qualified company managers due to the drive for efficiency and time pressure in operations. At the same time, there are also long-term strategic decisions where time pressure is no longer a factor, and yet cognitive bias appears, which has to be considered properly. In strategic issues, decision- makers tend to see their wishes and desires rather than the objective reality. The proposed system of fuzzy indicators based on technical and objective data supports decision-making between logistics strategies by mitigating cognitive biases, which is extremely important in the logistics field, where the decisions have to be made partly based on subjective, vague, or uncertain parameters.**

Index Terms—**logistics systems, cognitive bias, Push and Pull strategy, fuzzy description**

I. Introduction and Background

Cognitive infocommunication aims to create complex perceptual computing systems that effectively support human-machine communication [38]. The development of new methods, mathematical modeling, learning techniques, and related behavioral research will also help to better understand perceptual and cognitive brain processes [39]. For human-machine communication to be effective in the course of logistical strategies, it is necessary to identify and avoid cognitive biases in decision-making, so that its evaluation is not only based on intuition and subjective judgment alone. Data and pre-processing, standardization, and quantification are necessary to avoid associated biases.

A cognitive bias is a systematic deviation from rationality and logical, reasonable thinking and behavior. Cognitive biases are phenomena that influence thinking on experiences, intuition, and perceived things [1], thereby turning an objective decision into a subjective interest system [2].

Logistics, as a specialized field, typically requires decision-making at a daily level for the staff implementing logistics processes. On the other hand, logistics is often not seen as an independent field of expertise. As a consequence, the importance and necessity of the processes concerned are not assessed in sufficient depth, and decisions are based on intuition rather than rational decision-making. The time pressure characteristic of the logistics field [3] also significantly affects cognitive biases. Due to the particularities of logistics, there are departments where self-interest distortion [4,5] appears among cognitive biases.

People's thinking works together with simplifications, and cognitive biases [6]. For this reason, cognitive bias often occurs at the decision-making level in objective interest systems, which can be interpreted in several ways [7]. A systematic deviation from economic rationality in a company's decision-making model is called bias [8]. Psychologist Gary Klein [9], - who studied intuition in a scientific context at length and analyzed its effects in decision-making situations - named it a recognition-based decision model [9]. He concluded that intuition-based decisions only help managerial decisions in a predictable environment, similar to what has already been experienced countless times. So, decisions based on intuition are only acceptable if they are based on real experience [10].

The starting point for the appearance of cognitive bias is always a situation where a person responsible for decision-making receives information that they must incorporate into their decision-making mechanism [11]. They try to support the decision-making processes objectively, by examining facts and data [12]. As a result of unconscious prejudices, "beliefs", and expected results, the examination of facts and data becomes subjective, and decision-making takes place without self-checking [13].

Logistics decision-making takes place at different levels: it can distinguish between tactical, operational, and strategic decisions, which can be grouped according to a time horizon into short-, medium- and long-term decisions. In this paper, the 5 categories of cognitive biases will be described [14] and identified in the course of corporate decision-making and how and at what level these biases appear in the course of logistics decision-making will be shown.

The choice between Push and Pull systems are examined [15] as well as providing insight into the cognitive biases that appear during decision-making related to these systems. As it is not only determined by whether the production takes place to customer demand or stock, decision-making is often influenced by the cognitive biases that arise from fears about stock issues. First, where the names of the Push and Pull systems come from is mapped [16], and then the history of their development is presented [17]. After clarifying the concepts, it is explained which cognitive biases can appear during the logistics-related decision-making mechanism. Further, the limit will be examined at which a product can be reasonably defined to have

[1] Department of Logistics and Forwarding, Széchenyi István University, Győr, Hungary and Hungarian Research Network, Budapest, Hungary
[2] Department of Logistics and Forwarding, Széchenyi István University, Győr, Hungary. (E-mail: sos.eszter@ga.sze.hu)

DOI: 10.36244/ICJ.2024.5.13

a Push or Pull logistics strategy associated with it. For this purpose, a fuzzy measure is proposed that can be used to clearly identify the logistics strategy for the given product and the company's production.

The choice between Push and Pull logistics strategy for sustainability

One of the big trade-offs in Push-Pull is inventory versus delivery cost. This gives rise to logistical strategy decisions, during which cognitive biases may emerge as a significant deciding factor. The cognitive biases that emerge during the logistics decision-making process are introduced in Chapter 3, and the Push and Pull logistics strategies are discussed in detail in Chapter 4.

In this article, an analysis was carried out on the ScienceDirect and Scopus site on whether sustainability as a goal is reflected in the choice between Push and Pull logistics strategies. The aim of this publication is to present a methodology; therefore, the literature analysis is limited to these two databases.

An analysis was conducted on ScienceDirect and Scopus, first examining all publications, then narrowing it down to the last 5 years (2019-2024) and the 15 years before that (2004-2018), where the keywords in Table 1 and their contexts were included in the filtering: "push pull" AND (logistics OR "supply chain management") AND sustainability on ScienceDirect on the 10th of April 2024.

TABLE I.
NUMBER OF RELEVANT PAPERS. KEYWORDS: "PUSH PULL" AND (LOGISTICS OR "SUPPLY CHAIN MANAGEMENT") AND SUSTAINABILITY

PERIOD	KEYWORDS / RESULT SCIENCEDIRECT	NUMBER OF PUBLICATIONS	
		Science Direct	Scopus
2000-2024	Summary	326	1065
2019-2024	Last 5 years	159	834
2004-2018	15 years before the last 5 years	141	231

Table 1 shows that the search did not yield a large number of results, suggesting that there is currently little research on the relationship between logistics strategies and sustainability.

To examine the keywords of the publications, using the VosViewer software, which helps us analyze the relationships between keywords in the publications, showing the direction of the articles written in recent years in the context of Push and Pull logistics strategies and sustainability

Visualizing all the results of ScienceDirect (Appendix 1), it can be seen that the supply chain appears directly next to sustainability from 2018, and Industry 4.0 appears in 2021-2022, in addition to the circular economy, strategy, blockchain, technology adoption, and flexibility keywords. This led to the conclusion that professionals have recognized the need to develop long-term logistics strategies to achieve sustainability, yet the publications that have appeared have not focused on this aspect.

Based on the keywords, a visualization was made from the Scopus database, where the minimum number of occurrences was set to 20 due to the large number of keywords. In Appendix 2 it is already clear that from 2019-2020, terms related to environmental will also appear: environmental management, environmental technology, environmental regulations. Hence, environmental measures related to sustainability have become more prominent in the last 5 years. Appendix 2 also illustrates that in 2022, the keywords consumption behaviour and human will already appear, highlighting that more publications have already examined the human perspective in terms of logistics strategies and sustainability.

This paper aims to highlight the need to examine the Push and Pull systems and the cognitive biases in the choice of strategies in terms of sustainability, as the logistics strategy is fundamental to this.

II. LOGISTICAL ASPECTS OF COGNITIVE BIASES APPEARING IN CORPORATE DECISION-MAKING

In logistics, sustainability is primarily a strategic decision [18, 19]. Inappropriate logistics sub-processes damage the environment by purchasing unnecessary equipment, parts or packaging [20]. If logistics strategies are not internally coherent, this has an impact on sustainability. The presence of cognitive biases can be recognized in almost all areas of logistics. For their appearance to become clearly identifiable, cognitive biases are first described in the following.

Cognitive psychology basically defines two thinking systems. One system is characterized by conscious and processed thoughts [7], this is what is called "rational thinking" [21], which is not dealt with in this publication. The other thinking system is characterized by automatic and intuitive thoughts. The presence of experiences, prejudices, and assumptions causes cognitive biases in the system of thinking processes, thereby simplifying decision-making situations [22].

From the point of view of corporate decision-making, Olivier Sibony [14] classified cognitive biases into five main categories, within which 23 different types (see Appendix 3) were distinguished. Logistical aspects of cognitive biases appearing during corporate decision-making are as follows:

Pattern-recognition biases: Arise when a company tries to follow the example of a successful person by incorporating the same decisions into corporate strategies, but these decisions are not always appropriate for the company, which has a completely different corporate culture and product range [14].

Action-oriented biases: Refers to the cognitive biases that appear in the actions, which usually result from overly optimistic planning [23].

Cognitive inertia: When a process does not start due to certain facts and data. Loss aversion is one of the most powerful cognitive biases, which hinders change and the development opportunity that comes with it [24]. Loss aversion is of great importance in the subfields of logistics [25].

Self-interest biases: In some cases, for managers, the most important thing is not the company's lost money, but the loss of prestige resulting from their failure. Self-interest bias, also known as limited ethics, refers to cognitive biases that cause decent people to unknowingly show unethical behavior [4, 26].

Group biases: The cognitive biases created in the group mean that one individual's opinion has a significant influence on the decision-making structure of other people [27]. The effect of group bias is significant for the preparation of logistics strategies since logistics is often still negatively evaluated.

Based on the five categories, it can be clarified that the presence of cognitive biases in the decision-making mechanism related to logistics strategies can be definitely identified.

During the work of the persons responsible for the implementation of logistics processes in corporate decision-making, the logistics mindset often appears, casting logistics in a bad light, and causing in turn extra work and costs for the company. In practice, this means that the work of the people responsible for the implementation of logistics processes is treated as unwelcome and an unnecessary cost, so the managers of smaller companies tend to follow this way of thinking, both in terms of their work and their financial appreciation.

III. Concepts of Push and Pull strategies

The terms Push and Pull first appeared in Richard J. Schonberger's book in 1982 [16], comparing the Western-oriented "Push production system", which is based on the design philosophy of production resources and material resource planning; and the Japanese "Pull production system", which included the control technique based on Kanban logic together with the expectations of the Just In Time concept [36].

Nowadays, the definition of Push and Pull is already defined [15, 17], but professionals in the industry still often use the concepts of Push and Pull incorrectly, and consequently make bad strategic decisions. One of the big trade-offs in Push-Pull is inventory versus delivery cost. This gives rise to logistical strategy decisions, during which cognitive biases may emerge as a significant deciding factor.

In the case of a Push system, preliminary demand surveys are carried out, based on which the production program is prepared, the raw material is procured in the appropriate schedule, and then the production program is executed (Material Requirements Planning and Manufacturing Resource Planning MRP II).

In the case of Pull systems, production is always initiated by customer demand. In this way, the minimum stock level of the finished product can be ensured, but on the other hand, it means a longer lead time for the customer. The Just In Time (JIT) system enjoys great popularity among companies. However, its name causes a misunderstanding in the common language. From whose point of view is the product just in time? Undoubtedly from the point of view of the inventory as, due to the lead time, the customer always ends up waiting.

There is a trade-off in logistics between inventory or delivery (mobility). It connects to logistics strategies if the product is Push, more stock is needed and less transport, so you have less mobility, which means you have lower CO_2 emissions and less noise pollution. Conversely, if the Pull start strategy is investigated, there will be a higher number of deliveries associated with the freight transport, as the delivery of inventory is determined by customer demand. However, cognitive mobility [40] can lead to a deviation from what seems to be a good solution based on objective calculation.

The history of the development of Push and Pull systems dates back to the appearance of Material Requirements Planning [28], which enables the planning of material requirements for production and procurement [30]. The starting point was the number of final products defined in the production program, next the bill of materials was determined based on the material requirement, and then the gross component and raw material requirement [31].

The Material Requirements Planning (MRP) system was developed by Joseph Orlicky for the Toyota Manufacturing Program in 1964 [32]. Simultaneously, Black & Decker was the first company to use MRP. By 1975, MRP was implemented in 700 companies, and Joseph Orlicky's book Material Requirements Planning [32] was published in the same year. In 1983, Oliver Wight put the master schedule, rough capacity planning, capacity requirement planning, and other concepts into the classic MRP, thereby creating the basic idea of Manufacturing Resource Planning (that is MRP II) [33].

The Pull logistics strategy started with the Kanban system, which was introduced in the 1940s in supermarkets [32]. The order was determined based on the seller's inventory [15]. They only ordered more when the stock of the item was significantly reduced, thus optimizing the flow between the supermarket and the consumer. Toyota engineers noticed this method and, led by Taichi Ohno, investigated how it could be applied to work processes in the industry [33]. The Kanban initiates an action to replace the quantity consumed, so it is assigned to each production lot within the Just In Time (JIT) system [34,35]. To make Kanban effective, cycle time must be assigned [36].

Cognitive biases appear during decision-making between Push and Pull systems – case study

When a new product is introduced, the decision regarding the logistics strategy associated with it is generally made rationally and based on sound arguments. However, some concerns often lead to cognitive biases during the decision-making process related to Push and Pull strategies [17, 28], such as fear of the Bullwhip effect [29], lack of supplies, etc. An improperly chosen logistics strategy can cause supply disruptions and is also reflected in the company's processes. For example, the size of the warehouse is not only determined by how many products the company sells per month, but it is also significantly influenced by whether the manufactured products are made in a Push or Pull system. A product made entirely in the Pull system is not stocked, or only for a very short time. Products manufactured in the Push system are produced in large quantities [15], so they require a larger storage capacity.

Furthermore, it is a common problem for companies to use a Pull strategy until the managers are faced with the fact that the logistics processes are not working well. When investigating the reasons, it turns out that the Pull strategy was chosen because the manager who made the decision had previously worked for a company where the products were associated with the Pull logistics strategy. The decision is therefore accepted without any examination of the external circumstances. Considering Sibony's classification [14] (see Appendix 3) the presence of Overconfidence bias can already be clearly identified during decision-making (see Table.2).

TABLE II.
CROSS-CORRELATION OF BIASES (OWN COMPILATION BASED ON [14])

		Own knowledge	
		Underestimates	Overestimates
Position (power)	Underestimates	1 - Loss Aversion	2 - Experience bias
	Overestimates	3 - Halo effect	4 - Overconfidence

A cognitive bias appeared in the manager's decision-making mechanism [10], which, based on experience so far [11], led to the conclusion that the logistics strategy being implemented made the previous company successful, even though the product requires a completely different logistics environment. As a result of the appearance of the Halo effect [24], the appropriate logistics strategy was not implemented. The decision is partly based on existing work experience, so the presence of Experience bias [1] can be identified. After that, Status Quo [4] is a common phenomenon, when the management sticks to the originally formed decision. When it turns out that the right strategy has not been chosen for the given product, Loss aversion, Uncertainty aversion, and Hindsight bias [4; 14], which often arise during further decisions, appear as well.

IV. DEMARCATION OF PUSH AND PULL LOGISTICS STRATEGY

Push and Pull systems are determined based on the needs related to the product [15], but in many cases during production, it cannot be clarified whether a specific product is manufactured in a Push or Pull system. Fig. 1/a shows that the Completely Push System is when first the forecast (F) is made, the raw material is purchased, and only then does production begin, with the finished product then being sold by the vendors. In contrast, Fig. 1/b shows that only after the customer's order do the raw material procurement and production start. Then, when the order is fulfilled, the customer receives the product. In the Hybrid Push/Pull system shown in Fig. 1/c, the first half of the system behaves as a Push. The Material Decoupling point appears at the same time as the semi-finished stock, after which the product starts to behave as a Pull.

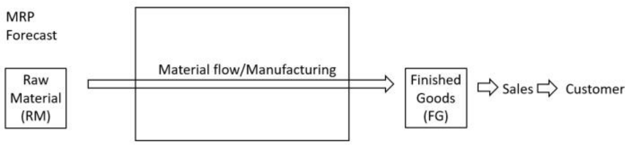


Fig.1/a. Completely Push System (MTS – Make-to-stock)

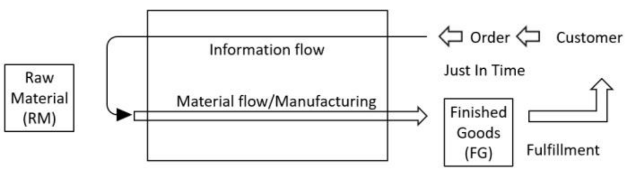


Fig.1/b. Completely Pull System (MTO – Make-to-order)

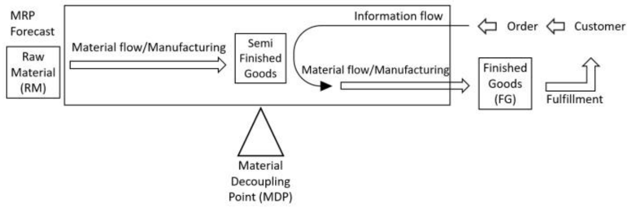


Fig.1/c. Hybrid Push/Pull System with Material Decoupling Point

V. FUZZY ASSESSMENT FOR THE DEMARCATION

An additional question is what characteristics can be used to determine whether it is Push or Pull. The first approach can be based on how many phases, for how long, and at what cost the material flow will proceed. For example, Fig. 2. shows a sequential production line

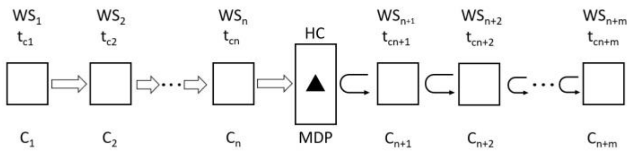


Fig.2. A sequential production line

Legend for sequential production line

Symbol	Meaning
WS_i =	The i^{th} Work Station (i= 1...n+m)
t_{ci} =	Cycle Time at WS_i. (i= 1...n+m)
C_i =	Production Cost at WS_i. (i= 1...n+m)
HC =	Holding Cost ($_{RM}$- raw material, $_{FG}$- finished goods)

The raw material arrives by WS_n, in the Push system. After that, it goes to the warehouse, and when an order comes, it will be pulled in small Kanban circles until WS_{n+m} (Fig.2). The MDP coincides with the entry into storage. There are $n+m$ workstations in total. Of these, n units are Push-based and m are Pull-based, the proportions of which are determined:

$$\text{Push: } \frac{n}{n+m} \tag{1}$$

$$\text{Pull: } \frac{m}{n+m} \tag{2}$$

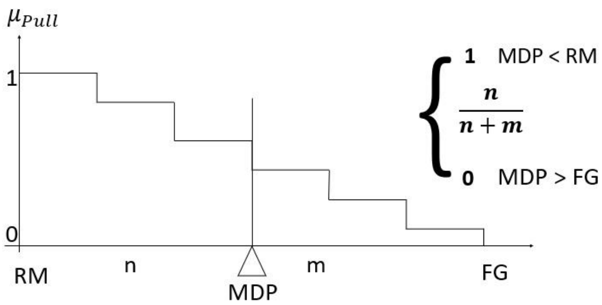


Fig.3. Fuzzy membership function of being pull calculated by (2)

This relationship considers only the used capacity in discrete numbers and can be used if the cycle times are nearly constant as shown in Fig. 3. For the representation of the actual status of the production concerning pull and/or push strategies fuzzy sets are introduced in which the membership values of being "Pull" are calculated at different accuracy, depending on what features of the production are taken into account (see Eq. (4), Eq. (6) and Eq. (8))

If the workload is not evenly distributed, it is necessary to consider the actual time spent in production (see Eq. (4)):

$$\text{Push: } \frac{\sum_{i=1}^{n} t_{ci}}{\sum_{i=1}^{n} t_{ci}+\sum_{j=n+1}^{n+m} t_{cj}} \quad (3)$$

$$\text{Pull: } \frac{\sum_{j=n+1}^{n+m} t_{cj}}{\sum_{i=1}^{n} t_{ci}+\sum_{j=n+1}^{n+m} t_{cj}} \quad (4)$$

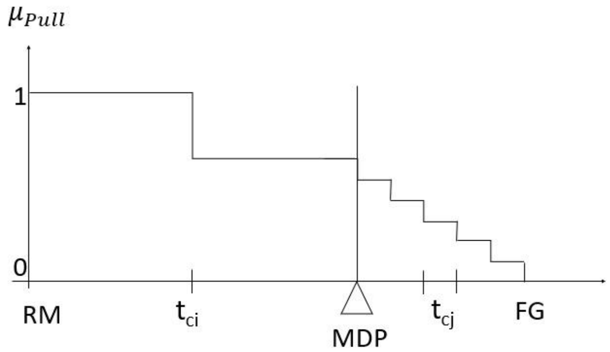


Fig.4. Fuzzy membership function of being pull calculated by (4)

By comparing Fig.3. and Fig.4. it is clear that the position of MDP alone is not enough to assess the “pullness” of the system. Without considering the cycle times judgement will be biased.

However, it does not yet include the complexity of the work process, for example, the processing may be different, such as rust protection, roughing, or fine machining. Hence a cost factor also must be included to determine the hourly price of the machine. In this case, C_i and C_j are introduced to represent the costs at each workstation (see Eq. (5) and Eq. (6)):

$$\text{Push: } \frac{\sum_{i=1}^{n} t_{ci} \cdot C_i}{\sum_{i=1}^{n} t_{ci} \cdot C_i+\sum_{j=n+1}^{n+m} t_{cj} \cdot C_j} \quad (5)$$

$$\text{Pull: } \frac{\sum_{j=n+1}^{n+m} t_{cj} \cdot C_j}{\sum_{i=1}^{n} t_{ci} \cdot C_i+\sum_{j=n+1}^{n+m} t_{cj} \cdot C_j} \quad (6)$$

C_i and C_j represent machine cost only, and there are further expenses to be considered. The cost of storage – in this case, Holding Cost *(HC)* -, is added to this (see Eq. (7) and Eq. (8)), where *RM* stands for raw material and *FG* stands for finished goods:

$$\text{Push: } \frac{HC+\sum_{i=1}^{n} t_{ci} \cdot C_i}{HC_{RM}+\sum_{i=1}^{n} t_{ci} \cdot C_i+\sum_{j=n+1}^{n+m} t_{cj} \cdot C_j} \quad (7)$$

$$\text{Pull: } \frac{\sum_{j=n+1}^{n+m} t_{cj} \cdot C_j}{HC_{FG}+\sum_{i=1}^{n} t_{ci} \cdot C_i+\sum_{j=n+1}^{n+m} t_{cj} \cdot C_j} \quad (8)$$

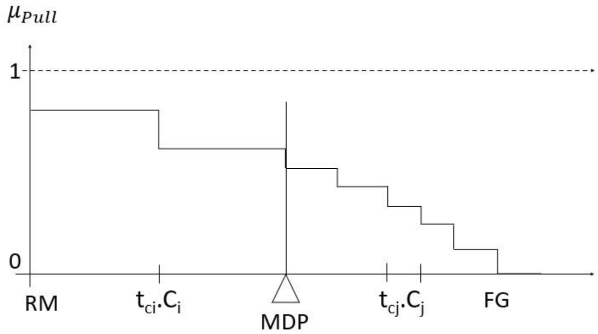


Fig.5. Fuzzy membership function of pull being calculated by (8)

If it contains HC_{RM}, it shows that the product is more of a Push product, since the Pull product is made for a specific period based on customer demand. So, the higher the *HC*, the more the system can be considered Push.

Important remark: Despite an apparently full pull system at first glance, when finished goods are stored for different reasons, for example, to consolidate the cargo and save cost in the delivery phase, in (Eq.8) HC_{FG} is not zero, the maximum value cannot reach *1*, and the sum of membership values of being push and being pull is also less than *1*. In this case, type2 fuzzy sets have to be applied that can describe the double uncertainty of the phenomena (see Fig. 5.).

The presented system of fuzzy representation is vital since - as it is demonstrated- even relatively simple situations can be very confusing and mislead the management. The proposed method can be used to objectively determine which logistic strategy is required for a given product, thus reducing the effect of Overconfidence bias in the process of decision-making.

The problem is well illustrated by the mask shortage that occurred during the COVID-19 pandemic. Mask production was a push system before the COVID-19 pandemic. A forecast was used to determine when and how many masks would be sold, and these forecast quantities were produced and then sold. With the onset of the COVID-19 pandemic, the forecast, based on decades of experience, was wiped out. Suddenly, production became a pull system, because when the raw material came in, they already knew which customer the particular quantity of masks they were producing belonged to. This publication demonstrates that the choice between Push and Pull systems is very confusing at times, as there are times when the Push logistics system is the right one, and then, in response to an unexpected situation, the product requires a Pull logistics system.

Accordingly, the system of equations described above can determine, for a given product or the company as a whole, whether the production uses a Push or Pull logistics strategy. Consequently, the decision-influencing effects of cognitive biases in the choice of logistics strategies can be avoided.

VI. Summary

Logistics is a frenetic field, one of the characteristics of which is frequent decision-making during the work process, accompanied all too often by time pressure. The importance of the logistics field was also brought to the fore by the COVID pandemic when most of society realized that if logistics systems do not work well, then the "supply of the world" does not work either. The disruption of supply chains also "overruled" the use of the previously popular Just In Time system, so the importance of choosing logistics strategies also increased.

This paper first presented an analysis of Push and Pull logistics systems and sustainability in terms of published papers. Despite the importance of addressing the impact of logistics strategies on sustainability, it was found that there are relatively few publications on the subject. After describing the emergence of cognitive biases, the 5 main categories of cognitive biases were introduced in corporate decision-making. For each main category, an example appearing in a logistics specialty was added to clarify the appearance of the types of corporate decision-making in the course of logistics decision-making. After describing the antecedents and history of the development of logistics strategies, the cognitive biases that appear during the choice between Push and Pull systems are illustrated through a concrete example. To clarify the choice between the Push and Pull strategy, three case studies were presented as well as a system of mathematical equations that support decision-making related to logistics strategies and help avoid the appearance of cognitive biases.

VII. Conclusion

This publication points out that the cognitive biases that appear during logistics decision-making have a significant impact on the formation of the logistics environment and logistics processes, so their influence on decisions cannot be ignored.

In the field of logistics, sustainability is first and foremost a strategic choice. Inappropriate logistics subprocesses damage the environment through the purchase of unnecessary equipment, parts, or packaging. If logistics strategies are not internally coherent, this, in turn, has an impact on sustainability. Cognitive biases occur during the decision-making process in the internal coherence of logistics strategies. Furthermore, the system may become unsustainable if an inappropriate logistic strategy is chosen because of cognitive bias. Therefore, it is also of paramount importance to avoid the appearance of cognitive biases in the decision-making process of logistics strategies to achieve sustainability.

The most important task of logistics strategy is to strengthen, coordinate and manage the relationship between corporate strategy and the logistics function, so that logistics can actively contribute to the company's performance and success. In this publication, only the Push and Pull logistics strategy was dealt with, which may seem simple, but if the right strategy is not chosen, it will not be sustainable.

Often Pull is chosen because it is fashionable and people think that keeping high stock is inappropriate. Assuming everything could be managed in Push and Pull, does not weigh the choice between strategies appropriately. However, if the Pull strategy is chosen unnecessarily, it can lead to high environmental impact, transport costs, pollution, and congestion in traffic areas.

It may be commercially worthwhile for a company because the external costs are not factored into the business model; society pays the cost, but it is not sustainable. In the short term, it seems right, but in the long term, it causes significant damage.

In order to give a proper description of the actual situation and help the decision-making, quantified measures have to be applied, and the fuzzy membership functions are capable of handling the issue. To develop the equation system, a sequential production line was first visualized, and then the workstations for the Push and Pull systems were defined. The membership values of the "Pull" state were calculated with different accuracies depending on the characteristics of the production. In order to estimate the "Pull" status, cycle times, workstation costs and storage costs had to be taken into account. The system of equations developed was used to convert the result into a type2 fuzzy number.

The proposed system of fuzzy indicators based on technical and objective data supports decision-making between logistics strategies by mitigating cognitive biases.

The presented and discussed set of fuzzy indicators based on technical and objective data are able to point out the real nature of the outlined production system from logistics point of view. The actual values of the calculated type2 fuzzy numbers are representation of the practical operation, so cognitive biases can be considered, which is extremely important in the logistics field, where the decisions have to be made partly on the basis of subjective, vague or uncertain parameters.

References

[1] Sibony O., Lovallo D., Powell T. C. Behavioral Strategy and the Strategic Decision Architecture of the Firm. *California Management Review*. 2017. Vol 59, Issue 3. DOI: 10.1177/0008125617712256

[2] Erceg N., Galić Z., Bubić A. Normative responding on cognitive bias tasks: Some evidence for a weak rationality factor that is mostly explained by numeracy and actively open-minded thinking. *Intelligence*. 2022. Volume 90, 101619 DOI: 10.1016/j.intell.2021.101619

[3] Albishri D. Y., Sundarakani B., Gomisek B. An empirical study of relationships between goal alignment, centralized decision-making, commitment to networking and supply chain effectiveness using structural equation modeling. *International Journal of Logistics Research and Applications. A Leading Journal of Supply Chain Management*. 2020. Volume 23, 2020 – Issue 4.

[4] Kahneman D., Tversky A. Prospect Theory: An Analysis of Decision Under Risk. *Econometric Society. 1979*. pp. 263–292. DOI: 10.2307/1914185

[5] Nadarevic L., Schnuerch M., Stegemann M. J. Judging fast and slow: The truth effect does not increase under time-pressure conditions. *Judgment and Decision Making*. 2021. Vol. 16, No. 5, pp. 1234–1266.

[6] Baron J. *Thinking and Deciding, 4th edition*. Cambridge University Press. 2006. ISBN: 9780511840265

[7] Toneatto T. Cognitive therapy for problem gambling. *Cognitive and Behavioral Practice*. 2002. Volume 9, Issue 3, pp. 191–199. DOI: 10.1016/S1077-7229(02)80049-9

[8] Stanovich K. E., West R. F. On the Relative Independence of Thinking Biases and Cognitive Ability. *Journal of Personality and Social Psychology 94*. 2008. pp. 672–695. DOI: 10.1037/0022-3514.94.4.672

[9] Klein G. *Sources of Power: How People Make Decisions*. Cambridge, MIT Press. 1998. ISBN13 9780262534291

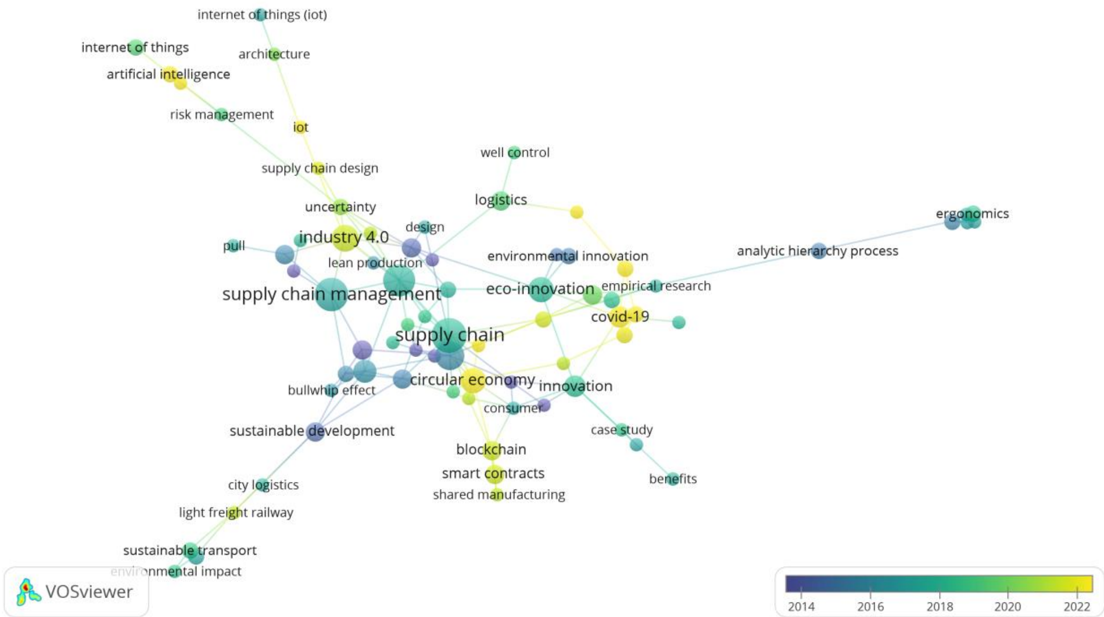
[10] Kahneman D., Klein G. Conditions for Intuitive Expertise: A Failure to Disagree. *American Psychologist 64*. 2009. pp. 515–526.

[11] Dobelli R. *The Art of Thinking Clearly*. Sceptre. 2013. ISBN: 9780062359803
[12] Kahneman D., Tversky A. Subjective probability: A judgment of representativeness. *Cognitive Psychology*. 1972. pp. 430–454.
[13] Trimmer P. C. Optimistic and realistic perspectives on cognitive biases. *Current Opinion in Behavioral Sciences*. 2016. pp. 37–43.
[14] Sibony O. *You're about to Make a Terrible Mistake: How Biases Distort Decision-Making and What You Can Do to Fight Them*. Little, Brown and Company. 2020. ISBN: 0316494984
[15] Puchkova A., Romancer J-L., McFarlane D. Balancing Push and Pull Strategies within the Production System. *IFAC-Papers OnLine*. 2016. Volume 49, Issue 2 pp. 66–71. DOI: 10.1016/j.ifacol.2016.03.012
[16] Schonberger R. J. *Japanese Manufacturing Techniques*. Macmillan, USA. 1982. ISBN: 9780029291009
[17] Minculete G., Olar P. “Push” and “Pull” systems in supply chain management. Correlative approaches in the military field. *Journal of Defense Resources Management*. 2016. Vol. 7, No. 2. pp. 165–172. ISSN: 2068-9403
[18] Golam K., Sanjoy K. P. and Syed M. A. (2022) Sustainable Assessment in Supply Chain and Infrastructure Management. *Sustainability. 14(11)*, 6787. DOI: 10.3390/su14116787
[19] Rajeev A., Rupesh K. Pati, Sidhartha S. Padhi, Kannan, G. Evolution of sustainability in supply chain management: A literature review. *Journal of Cleaner Production*. Vol. 162, pp. 299–314. DOI: 10.1016/j.jclepro.2017.05.026
[20] Sós E., Földesi P. Fuzzy QFD assessment of logistics coherence. *2022 IEEE 6th International Conference on Logistics Operations Management (GOL)*. 2022. pp. 1–6. DOI: 10.1109/GOL53975.2022.9820009
[21] Waerneryd K. E. *Stock-market psychology*. Northampton, Edward Elgar Publishing. 2001. ISBN: 978 1 84064 736 5
[22] Kahneman D. Thinking, *Fast and Slow*. Macmillan Publishers. 2011. ISBN: 9781429969352
[23] Moore D. A., Healy P. J. The Trouble with Overconfidence. *Psychological Review 115*. 2008. pp. 5002–5017. DOI: 10.1037/0033-295X.115.2.502
[24] Rosenzweig P. *The Halo Effect... and the Eight Other Business Delusions That Deceive Managers*. New York, Free Press. 2007. ISBN: 1471137163
[25] Földesi P., Botzheim J.: Modeling of loss aversion in solving fuzzy road transport traveling salesman problem using eugenic bacterial memetic algorithm. *Memetic Computing 2*. 2010. pp. 259–271.
[26] Janis I. L. Groupthink: *Psychological Studies of Policy Decision and Fiascoes*. Boston, Wadsworth. 1982. ISBN-10: 9780395317044
[27] Bazerman M. H., Loewenstein G., Moore D. A. Why Good Accountants Do Bad Audits. *Harvard Business Review 80(11)*. 2008. pp. 96–102, 134.
[28] Puchkova A., Romancer J. L., McFarlane D.: Balancing Push and Pull Strategies within the Production System. *IFAC-Papers OnLine*. 2016. Volume 49, Issue 2, pp. 66–71. DOI: 10.1016/j.ifacol.2016.03.012
[29] Lin J., Zhou L., Spiegler V. L. M., Naim M. M., Syntetos A.: Push or Pull? The impact of ordering policy choice on the dynamics of a hybrid closed-loop supply chain. *European Journal of Operational Research*. 2022. Volume 300, Issue 1. pp. 282–295. DOI: 10.1016/j.ejor.2021.10.031
[30] Buzacott J. A.: Queueing models of Kanban and MRP controlled production systems. *Engineering Costs and Production Economics*. 1989. Volume 17, Issues 1–4, pp. 3–20. DOI: 10.1016/0167-188X(89)90050-5
[31] Milne R. J., Mahapatra S., Wang C-T. Optimizing planned lead times for enhancing performance of MRP systems. *International Journal of Production Economics*. 2015. Volume 167, pp. 220–231. DOI: 10.1016/j.ijpe.2015.05.013
[32] Kiran D.R. Chapter 30 – Material requirement planning. *Production Planning and Control, A Comprehensive Approach*. 2019. pp. 429–439. DOI: 10.1016/B978-0-12-818364-9.00030-5
[33] Kiran D. R. Chapter 31 – Manufacturing resource planning (MRP II). *Production Planning and Control, A Comprehensive Approach*. 2019. pp. 441–455. DOI: 10.1016/B978-0-12-818364-9.00031-7
[34] Sudarsan D. and Sivakumar G. D. Implementation of kanban to improve efficiency of a manufacturing system subject to bottlenecks. *International Journal of Manufacturing Technology and Management*. 2020. pp. 467–480.
[35] Kozma D. and Varga P. Supporting Digital Supply Chains by IoT Frameworks: Collaboration, Control, Combination. *Infocommunications Journal*, Vol. XII, No 4, December 2020, pp. 22–32. DOI: 10.36244/ICJ.2020.4.4
[36] Kiran D.R. Chapter 26 – Just in time and kanban. *Production Planning and Control, A Comprehensive Approach*. 2019. pp. 369–379. DOI: 10.1016/B978-0-12-818364-9.00026-3
[37] Olhager J. The role of the customer order decoupling point in production and supply chain management. *Computers in Industry*. 2010. Volume 61, Issue 9, pp. 863–868.
[38] Baranyi P., Csapo A., & Sallai G. (2015). *Cognitive infocommunications (coginfocom)*. Springer Book
[39] Baranyi, P., & Csapó, Á. (2012). Definition and synergies of cognitive infocommunications. *Acta Polytechnica Hungarica*, 9(1), 67–83.
[40] Zöldy, M. & 39. Baranyi, P. (2023) The Cognitive Mobility Concept. *Infocommunications Journal*, 35–40.

Péter Földesi PhD was born in Budapest, Hungary in 1962. He graduated from the Budapest University of Technology and Economics as a Transport Engineer in 1986 and obtained PhD degree at the Hungarian Academy of Sciences in 1994. He worked for a large international haulage company as a marketing consultant in 1990, then joined the academic staff of Budapest University of Technology and Economics. In 1995 he moved to Győr (Hungary). He took a position at the Széchenyi István University, at the Department of Logistics and Forwarding, where he was promoted to Head of the Department in 2007. He was the Rector of the University between 2013 and 2021.

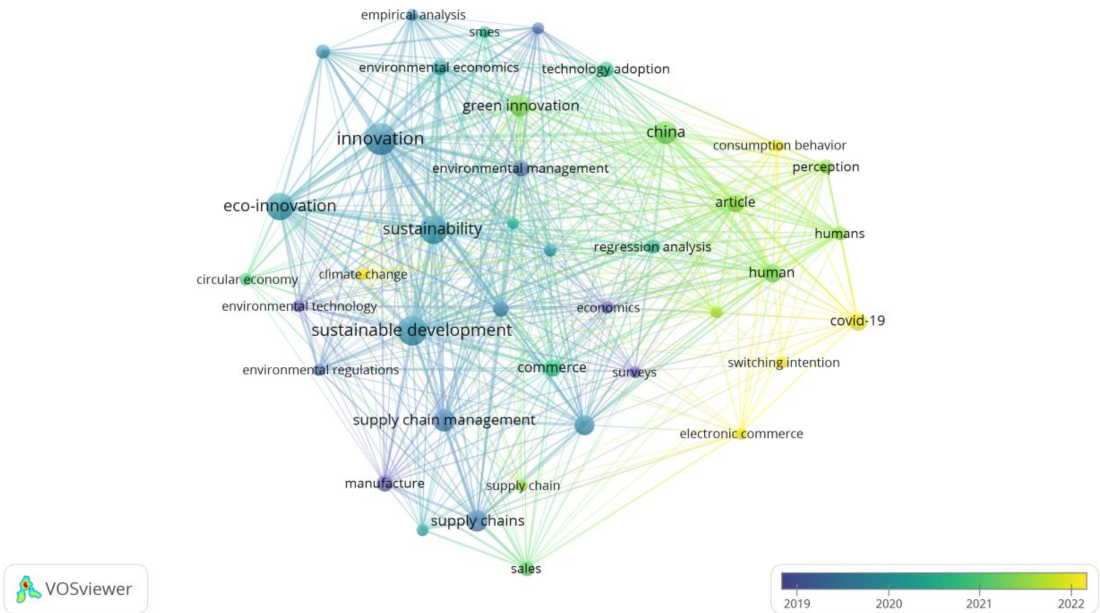
Eszter Sós was born in Bonyhád, Hungary in 1984. She graduated as a Logistics Engineer and is pursuing her PhD studies at the Doctoral School of Multidisciplinary Engineering Sciences at Széchenyi István University. She worked in the industry and logistics services for several years before starting her academic life. She is an Assistant Lecturer at Széchenyi István University, in the Department of Logistics and Forwarding. Her research focuses on the structured design of logistics strategies, including the internal coherence of logistics strategies. Furthermore, she investigates the emergence of cognitive biases in decision-making processes in the logistics field.

APPENDIX 1



OVERLAY VISUALIZATION BY VOSVIEWER FROM THE RESEARCH RESULTS OF SCIENCEDIRECT ON THE 10TH OF APRIL 2024. KEYWORDS: "PUSH PULL" AND (LOGISTICS OR "SUPPLY CHAIN MANAGEMENT") AND SUSTAINABILITY. THE MINIMUM NUMBER OF OCCURRENCES OF A KEYWORD: 2.

APPENDIX 2



OVERLAY VISUALIZATION BY VOSVIEWER FROM THE RESEARCH RESULTS OF SCOPUS ON THE 10TH OF APRIL 2024. KEYWORDS: "PUSH PULL" AND (LOGISTICS OR "SUPPLY CHAIN MANAGEMENT") AND SUSTAINABILITY. THE MINIMUM NUMBER OF OCCURRENCES OF A KEYWORD: 20.

APPENDIX 3

Types of cognitive biases	Pattern recognition biases	Storytelling trap	Instinctively searching for information that provides justification to support a situation.
		Confirmation bias	Searching for corroborating information about pre-existing views, e.g. policy. Ignoring information that contradicts our point of view.
		Champion bias	The company trusts a well-performing professional more, even if the information is objectively not relevant.
		Experience bias	To make a decision based on memory and experience in an environment where these experiences are no longer applicable.
		Attribution error	Success (or failure) is attributed to a single person in the company. E.g. Steve Jobs, Apple
		Halo effect	Adopting the best practices of a successful person in the hope that the strategy they use will work.
		Survivorship bias	Only the strategies of successful leaders are studied. There are several lessons to be learned from studying failed companies.
	Action-oriented biases	Overconfidence	To overestimate one's abilities, which affects the decisions made.
		Planning fallacy	Excessively optimistic planning, regarding budget and time.
		Overprecision	To overestimate our ability to predict the future.
		Competitor neglect	Underestimating competitors, ignoring the competition
	Cognitive inertia	Anchoring	The leader tends to use the numbers presented to him as an "anchor" even if that number is not relevant to the case. For example, an annual budget plan.
		Commitment escalation	Because of a promise, they won't change the already established strategy so that the energy invested until then "doesn't go to waste"
		Status quo bias	Avoiding difficult decisions, therefore maintaining the status quo.
		Loss aversion	Loss aversion is one of the biggest barriers to development.
		Uncertainly aversion	Avoiding unknown risk in order to avoid loss.
		Hindsight bias	They judge the same occurrence with different probability before and after the event has occurred.
	Self-Interest biases	Present bias	The company's management does not think long-term. Immediate benefits instead of future profits.
		Self-serving bias	Individuals are driven by an unconscious intention to make decisions for their own self-interest, whether financial or emotional.
		Omission bias	The management of the company accepts when someone else makes a mistake and not themselves. In addition, they consider it morally acceptable to profit from it.
	Group biases	Groupthink	People tend to adopt the collective point of view even when they know it is not correct.
		Information cascades	The order of speakers distorts the outcome, as the opinions of those who speak first are amplified.
		Polarization	As a result of groupthink, the group's opinion will be more extreme, which the group members will represent more confidently. It also deepens commitment.

Design Suggestions for Digital Workflow Oriented Desktop VR Spaces

Anna Sudár[1,2,3,*], and Ildikó Horváth[1,2,3,*]

Abstract—**This paper presents an examination of design principles for 3D Virtual Reality (VR) environments, with a focus on enhancing digital workflows. Employing objective data, the study sets out to clarify the primary design considerations for crafting effective 3D VR spaces. Through empirical research, the authors conducted comparative analyses of task performance within both classical 2D Windows and in 3D VR environmental contexts, exploring users' perceived difficulty levels alongside eye-tracking data. The findings reveal that, although 3D VR environments rich in distracting elements and demanding high navigational effort increase perceived task difficulty, these factors do not negatively impact overall performance or task completion time. Interestingly, eye-fixation duration results indicate that visual fixation in 3D VR falls within expected norms, whereas in 2D scenarios, fixation rates are significantly higher, more than doubling those observed in 3D settings. Drawing on these insights, the paper supports the design of 3D VR spaces that are simpler and intuitive, necessitating minimal navigation, thereby optimizing task performance efficiency.**

Index Terms—**virtual reality, desktop VR, virtual space design**

I. Introduction

Ergonomic considerations in Virtual Reality (VR) are crucial for ensuring a comfortable, safe, and engaging user experience and supporting performance efficiency. These considerations primarily address the physical interaction between the user and the VR environment to minimize discomfort and potential health risks. A large volume of scientific studies support the idea that user experience (UX) design is crucial, as good UX design has been shown to enhance engagement and motivation, and can help maintain user attention for longer durations compared to traditional 2D interfaces [1]–[3]. These effects also extend to 3D spaces, as they allow users to create, visualize, and recall information in visually appealing and persuasive learning environments [4], [5]. Desktop VR combines elements of traditional computer use with the immersive qualities of VR, making it essential to consider both digital ergonomics and the unique demands of a 3D virtual environment. In today's digital interface design, which is connected to task performance and knowledge acquisition, it is advisable to take into account research results based on eye-tracking technology measurements. These studies have found correlations between fixation numbers and task difficulty, and fixation durations might be influenced by underlying affective processes that contribute to learning. [6]–[9].

[1] Corvinus Institute for Advanced Studies, Corvinus University of Budapest, Hungary
[2] Institute of Data Analytics and Information Systems, Corvinus University of Budapest, Hungary
[3] Hungarian Research Network, HUN-REN, Budapest, Hungary
(*Email: {anna.sudar, ildiko.horvath}@uni-corvinus.hu)

DOI: 10.36244/ICJ.2024.5.14

The UI should be intuitive and easy to navigate. One primary consideration in designing spatial elements is to take into account human spatial perceptions. Accordingly, it is recommended to design virtual spaces as open areas or, in the case of closed spaces, with high ceilings [10]. Good quality audio that is synchronized with the visual elements of the VR environment can enhance the experience without causing auditory strain. Text and icons need to be large and clear enough to be easily readable. Interaction methods should be natural [11] and not require excessive or uncomfortable movements or complicated workflows. The ability for users to customize their VR experience (like adjusting sensitivity settings, UI elements, etc.) can greatly improve comfort and accessibility for a wide range of users. In this study, drawing from our research results, we aim to establish design guidelines aimed at enhancing user productivity and working effectiveness and contributing to the advancement of desktop VR environments. The structure of the paper is organized as follows: In the second section, the Definition and Metrics are presented. This section introduces key concepts, definitions, and the metrics used to evaluate virtual reality environments. The next section is the intersection of Cognitive Infocommunications (CogInfoCom) [12], [13] and cognitive aspects of VR (cVR) [14] technologies, presenting relevant research findings. The Research Context sets the background for understanding the framework of this study. Then, the authors present the main finding of the study including data analysis, interpretations, and the implications of research within the VR domain. The session Discussion critically examines the research results, discussing their significance, the limitations encountered, and how they relate to existing literature. Based on this research, the authors propose actionable recommendations for designing immersive and user-friendly VR environments to support performance efficiency. The paper concludes by summarizing the key findings, their implications for the field of VR, and suggesting directions for future research.

II. Definition and metrics

The objective of this chapter is to define essential terminology, providing clarity and ensuring a comprehensive understanding of the terms employed throughout the paper. Several of the key definitions presented herein have their origins in prior publications by the authors. However, it is imperative to emphasize that their application within this context is entirely novel and original.

A. Digital workflow—DW

[15] Digital workflows determine the order in which individual digital elements are to be accessed or processed during the course of a digital project. We distinguish among the following types of digital workflows:

1) 1st order (linear): The digital elements are to be accessed in a static and sequential order, one after the other.

2) 2nd order (loopy): There are loops in the order in which the digital elements are to be accessed, so that individual elements, or smaller sequences thereof, are to be accessed repetitively. Such loops can be characterized by length and number of repetitions.

3) 3rd order (networked): Digital elements accessed during the project are structured as hierarchical loops, so that the project may contain subprojects of subprojects, and/or the ordering of digital elements may be different upon different repetitions of the loops.

4) 4th order (algorithmic): It is possible that the project contains branches, so that different digital elements are accessed dynamically in an order that depends on information obtained during the project.

B. Digital Guidance—DG

[15] Digital guidance is taken to mean a process that unambiguously drives the user's attention during the digital workflow and thus reduces (partially, or to 0) the time required for searching for and finding the relevant digital content. It is possible to distinguish among three forms of digital guidance as follows:

1) none: no guidance is applicable, or the representation of the digital content doesn't involve embedded digital elements (instead, the elements are provided through separate lists).

2) sequential (DG-S): The digital elements are traversed in sequential order. It is thus possible to jump from one element to the next in the context of a digital workflow.

3) random access (DG-R - event/dynamic focus-driven): One can switch between sequences of digital elements, and thus follow non-static sequences (for example, in the case of DWs of the 4th order).

C. Information Availability—IA

[11] This indicates what percentage of the information (digital content) needed to execute a workflow is available in the digital work environment when executing up to 1 Navigation Based elementary Operation.

Remark: e.g., 100 percent, if all the information required to execute the workflow is available and can be accessed by a navigation operation.

D. Information Access Cost—IAC

[11] The weighted sum of the time spent accessing information for each type of operation, where the weights are the number of elementary operations corresponding to that type of operation.

$$IAC = \sum_{i=1}^{n} O_i * t_i \qquad (1)$$

Remark: this metric is high even when complex operations are performed in the same amount of time and when simple operations are performed for a long time. This includes the complexity of the operation and the user's ability.

E. Information Validity—IV

[11] This indicates the percentage of the information presented in the digital work environment that is directly required for task execution.

IV = Number of Valid Information Units/Summa necessary information units number.

F. Personalized Workflow Order Ability—PWO

[11] This is an indicator of the facility of the digital work environment to provide users with the ability to arrange digital content in their own way. The value is 0 if the option is not provided, 1 if the users can set the layout themselves, and 2 if the optimal layout is automated with AI support.

G. Personalized Information Overview—PIO

Indicates the ability of the digital work environment to provide users with the ability to set personalized information overview. The value is 0 if the option is not given, 1 if the users can set the layout themselves, and 2 if it is automated with AI support.

H. Preference point

Preference points in the virtual space that users were more likely to visit in order to find the spots that best allowed them to oversee the space and solve their tasks.

I. Content arrangement types

Patterns with which users preferred to arrange the content

1) Content: content types that had a similar subject matter were most likely to be arranged in clusters, in close proximity to each other.

2) Type: similar content types were most likely to be arranged in clusters, close to each other.

3) Mixed: primal organizing principle is the content, but within the same area, the secondary organizing principle is the type.

III. Antecedent researches by CogInfoCom and cVR

CogInfoCom stands at the intersection of infocommunications and cognitive sciences, striving to enhance the co-evolution and interplay between human cognition and digital technology. This interdisciplinary field is dedicated to elevating human efficiency in digital workspaces and refining work processes through innovative IT solutions. It delves into the mutual evolution of infocommunication devices and human cognitive functions, aiming to optimize interaction within digital environments [12], [13], [16].

cVR, on the other hand, delves into the expansion of human cognitive abilities through the utilization of various technologies within a three-dimensional spatial framework.

Here, VR systems, as advanced infocommunication devices, play a pivotal role. They revolutionize information organization and management, allowing for spatial categorization and prioritization of data according to its significance in the workflow. VR's capacity for collaborative use and integration of other technological advancements further underscores its value in this domain [14].

CogInfoCom and the cVR field are extensively researched scientific domains, showcasing significant advancements such as the development and enhancement of the three-dimensional virtual library model [17], [18]. This model opens up new possibilities in digital information architecture. The fields also delve into 'mathability,' which investigates the synergy of artificial and natural cognitive abilities in mathematics [19], [20]. Furthermore, they are at the forefront of pioneering new educational methodologies, or foreign language education and linguistics [21], and are actively exploring learning challenges and emerging opportunities in the rehabilitation of autistic children [22], underlining their expansive and dynamic nature.

Moving to the connection with UI design, VR systems herald a paradigm shift from conventional command-based interactions to dynamic, user-centric interfaces. This shift necessitates a deep understanding of human behaviour within these virtual environments, provided by cognitive science. The inherent properties of VR, coupled with insights into human cognition, pave the way for more sophisticated and efficient virtual workspaces compared to traditional 2D interfaces. The introduction of a third dimension not only facilitates more organized information management but also resonates with the innate spatial understanding of human users. This spatial familiarity, along with the use of metaphors, contributes to a robust comprehension of tasks and data. Consequently, desktop VR systems enable collective visualization and observation, fostering an environment conducive to the sharing of knowledge and information [23].

VR brings new challenges and opportunities in the context of UI and information management. While traditional Windows interfaces offer users a familiar, structured environment controlled by mouse and keyboard, VR is a digital interface that provides interaction in a much more immersive spatial environment and operates according to a completely different paradigm. In recent years, several publications in the field of CogInfoCom and cVR have presented research results related to VR design. Of these, the relevant results for this study are:

Virtual Reality (VR) introduces a dynamic array of challenges and possibilities within the realms of User Interface (UI) design and information management. Contrasting the conventional, well-known setup of Windows interfaces, which users navigate via keyboard and mouse, VR offers a novel and immersive spatial interface operating under a fundamentally different paradigm. Recent years have seen a surge in scholarly work within Cognitive Infocommunications (CogInfoCom) and Cognitive Aspects of Virtual Reality (cVR), contributing significant insights into VR design.

Experimental research has shown that preference points and attentional focus points can be found in 3D virtual reality spaces. In three-dimensional virtual reality settings, users tend to concentrate more and spend additional time at certain points of interest, which helps them navigate and complete assigned tasks [24]. Moreover, using virtual reality interfaces has been linked to improved recall of the process involved in organizing three-dimensional objects in a given space [25], [26], and also enhances the retention of information distributed across a space [25]. Cannavò and colleagues suggest that automating the conversion from two-dimensional to three-dimensional formats can significantly enrich the user experience in virtual reality workspaces, making it more engaging and productive [27].

Berki and associates conducted a study comparing how well users remember images shown in a three-dimensional virtual reality setting as opposed to a two-dimensional website. They discovered that the virtual reality environment was more effective in aiding memory recall [28], [29]. In a similar vein, it was observed that desktop virtual reality systems outperform traditional two-dimensional browsers in memory retention of additional information [30].

In previous research conducted by one of the contributors to this study, the effectiveness of personalization based on learning styles in three-dimensional desktop virtual reality was examined. The findings from that research revealed that when the instructional content was tailored to their individual learning styles, users achieved scores that were 20 percent higher, along with an 8–10 percent improvement in response times in subsequent assessments [31]. This customization of the learning environment demonstrated a significant impact on user behaviour and performance in the 3D virtual reality context.

These studies collectively indicate that virtual reality technology may play a crucial role in reducing the mental effort required for spatial memory and in improving performance in tasks that involve spatial orientation, particularly those that involve navigating through a vast array of digital documents.

In virtual spaces, the creation of a sense of presence is of utmost importance, and it is equally crucial to bestow upon the user a feeling of control [32]–[34]. This serves several purposes. Firstly, in reality, individuals have control over their own bodies, primarily altering their positions, orientations, and perceptions of their environment [35], [36]. This agency must be replicated within virtual worlds to ensure the user's sense of security. Additionally, research has shown that active participation and control over one's environment lead to better retention of events and information [37], [38]. Thus, providing users with the ability to actively shape their virtual surroundings enhances their sense of control.

The concept of control is closely tied to perception, which, in the real world, flows through various modalities such as visual, auditory, and haptic channels. These modalities aid in navigation, information processing, storage, and even survival [39]. In the realm of virtual reality, almost all these modalities can be simulated. However, haptic feedback remains less developed in desktop VR, lagging behind the tactile sensations provided by controllers, gloves, and similar devices. Additionally, replicating olfaction, the sense of smell is currently a limitation in virtual technology. Nevertheless, groundwork has been laid to address these challenges and further enhance sensory immersion in virtual environments [40]–[42].

Special Issue
of the Infocommunication Journal

IV. Research background

A. Subjects

A total of 21 participants were divided into two groups for each condition, and no specific qualifications or expectations were considered when selecting participants. For the 3D VR measurement, 14 participants initially took part, but due to technical issues and breaks in data collection, the results from nine subjects (three women and six men, aged between 17–55 years, with a mean age of 32.5 years) will be published in the analysis. The 2D environment study involved seven participants (four women and three men, aged between 25–33 years, with a mean age of 27.83 years). All participants were Hungarian native speakers, volunteering for the experiment with informed consent obtained in advance. The research was conducted under institutional endorsement ensuring ethical compliance and data privacy.

B. Procedure

This study explores some of the cognitive capabilities of users in desktop VR environments by comparing spatial behaviours, performance, and subjective experiences between traditional 2D interfaces and immersive 3D VR settings. At the initiation of the session, participants engaged in the prescribed task within three dimensions were requested to demonstrate their familiarity with the MaxWhere software and provide an estimation of the time they had allocated to using the application. Those participants who lacked prior knowledge of the software practised for around 30 minutes in acquainting themselves with its functionalities and acquiring fundamental user capabilities. Proficiency in essential skills encompassed adept navigation within the software, as well as the ability to activate and deactivate the display panels and engage with the content presented on them. The evaluators assessed the mastery of proficient software utilization. Participants were tasked with studying materials related to astronomy and completing questionnaires in either a 2D environment hosted on Google Sites or a 3D environment using the MaxWhere desktop VR platform. Despite using identical study materials and questionnaires in both settings, they were presented differently: linearly in the 2D interface and spatially in the 3D VR environment as shown in Figure 1. The questionnaires that were specifically designed to align with each of the subtopics were composed of a series of true-or-false questions, a set of multiple-choice questions, as well as a collection of questions that necessitated brief responses typically consisting of only one or two words. The task was a reading comprehension exercise which is very common in education and also in work scenarios, where participants found all the necessary information in the digital materials placed around the questionnaires. These materials were essential for answering the questions posed in the questionnaires. One example of the true-or-false question type: ”Black holes can be observed through their gravitational effects on the surrounding gas, dust, and stars. True or False?” - The participants were asked to find the information around the questionnaire in on of the related digital contents and mark in the questionnaire the right answer.

The theme of all questionnaires revolved around astronomy. The four subthemes of the questionnaires were (“*Universe*”, “*Planets*”, “*Satellites*” and “*Space Research*”). We chose this topic because none of the participants were experts in the field, ensuring that they approached the tasks with similar levels of knowledge. This prevented any significant disparities in results caused by someone being highly knowledgeable in certain areas and potentially skewing the measurement outcomes.

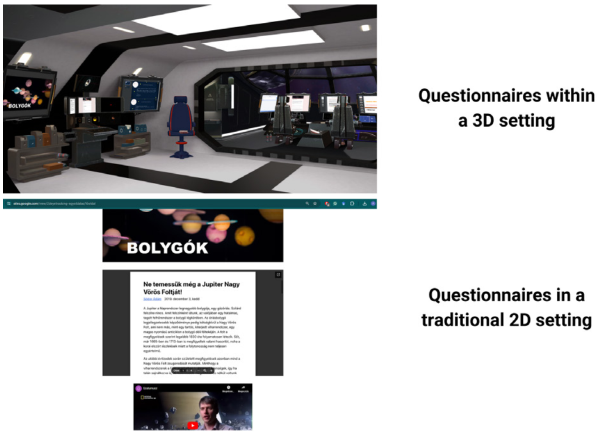


Fig. 1. This Figure illustrates through an example that the same digital content and questionnaires appeared for participants in both the 2D and 3D scenarios. The difference was visibly apparent in the layout mode.

C. Technical background

We used an HP Omen 15 laptop for conducting the research, which had the following specifications:

- AMD Ryzen™ 7 5800H processor
- NVIDIA® GeForce RTX™ 3060 graphics card
- 15.6-inch screen
- 16 GB RAM
- 512 GB SSD

The device used for Eye Tracking measurements was an EyeTribe eye tracker along with OGAMA (Open Gaze and Mouse Analyzer) Version 5.1 software. Since the VR software used for the measurement was a desktop VR application, it did not require the use of extra peripherals (such as an HMD). The eye-tracker was positioned at the bottom of the screen at an appropriate angle, ensuring accurate and consistent data collection for each individual.

D. Experimental environment design

When designing the virtual environment, special attention was paid to incorporating the experiences gained from previous research into the 3D virtual space created for this measurement. Ensuring dynamic interaction with the content was an important consideration because we assumed that such a setup would better meet the demand for a holistic overview of the educational materials, similar to placing documents on a physical desk. This could potentially lead to more effective learning outcomes and more effective performance compared to the more static and linear presentation of materials in the 2D environment.

The experiment was designed with several key objectives structured around both quantitative and qualitative assessments:

Objective Performance Metrics, This includes quantifiable data such as:

- The total number of correct answers participants provided.
- The amount of time participants spent completing the questionnaires.

Subjective Assessment: This assessment was based on the participants' personal evaluations of:

- The difficulty level of the questionnaires.
- Interaction Patterns in a 3D Environment: The study also examined how participants interacted within a three-dimensional (3D) environment, focusing on specific behaviours and engagement methods.

V. Research Results

A Mann-Whitney U test was used to compare the scores of the 3D Group (Mdn = 614.07) and the 2D Group (Mdn = 317.16) (descriptive statistic in Fig. 1) on the fixation duration mean. The Mann-Whitney U statistic was U = 55, indicating a significant difference between the two groups (p = 0.012) (Fig. 2).

	Fixation Duration Mean (ms)	
	2D	3D
Valid	7	9
Missing	0	0
Mean	766.390	311.948
Std. Deviation	558.407	42.459
Minimum	289.074	254.968
Maximum	1916.796	392.029

Fig. 2. Descriptive statistics of the fixation duration mean between the two measured groups.

We conducted an independent samples t-test to assess the differences between 2D and 3D environments across various thematic sections. Each analysis was preceded by a Levene Test for Equality of Variances, confirming homogeneity in variances across comparisons. The results indicated no significant differences between the groups, except in the time required to complete the satellite-themed questionnaire. Participants in the 3D environment completed this task noticeably quicker (units of measure were minutes), showing a significant difference ($t(14) = 3.38$, $p < 0.05$).

In the 3D virtual space, participants found the sections on Satellites and the Universe the most challenging, with an average score of 4.25 out of 7 (SD: 1.6). This was followed by the Planets section with a score of 3.75 (SD: 1.42), and

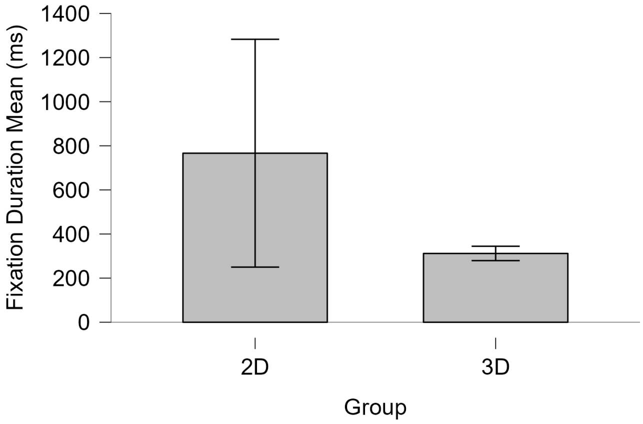


Fig. 3. Fixation duration mean differences between the 3D VR and the 2D group.

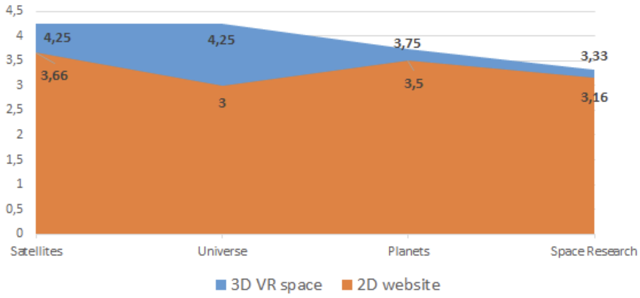


Fig. 4. Subjective user feedbacks - Comparative difficulty and performance in 2D vs. 3D virtual learning environment.

the Space Research section was the easiest at 3.33 points (SD: 2.01). Conversely, participants navigating the 2D website rated the Satellites section as the most difficult, scoring an average of 3.66 (SD: 1.63), followed by Planets at 3.5 (SD: 1.22), Space Research at 3.16 (SD: 1.94), and the Universe section as the easiest with a mean score of 3 (SD: 1.095). Fig 3. shows the research results connecting to these user feedbacks.

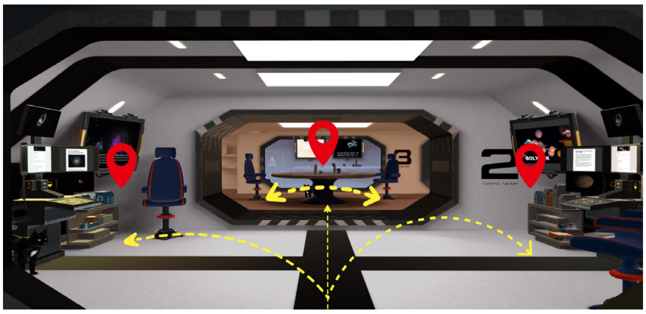


Fig. 5. Interaction routes and overview point examples.

Based on the analysis of screen recordings, a prevalent interaction patterns (Figure 5.) observed among the participants in the 3D measurement was what we refer to as the "overview mode". The red marks in Figure 5. indicate some examples

Recommendation	Description	Relevant research results
Improve Multimodal Stimulus and Feedback	Use auditory and visual contents and feedback to enhancing immersion and information delivery.	- Subjective assessment of the difficulty level of the questionnaires - 2D 3D differences and content analysis
Ensure Consistency	Apply consistent UI/UX design principles, accommodating dynamic VR interactions and perspectives.	- Information overview - Fixation duration time results
Minimize Textual Instructions	Prioritize visual and spatial cues over textual content for guidance, supporting seamless user engagement.	- Subjective assessment of the difficulty level of the questionnaires - Fixation duration time
Incorporate Familiar Elements	Utilize familiar interfaces and interaction models from 2D environments to improve VR usability and ease the transition.	- Overview interaction - Fixation duration time.
Strategic Content Placement	Organize digital content for easy navigation and task management, using circular arrangements and prominent viewpoints.	- Overview, alternating view.
Optimize Object and Content Placement	Consider the size, type, and spatial relationships of content and objects to support tasks without excessive cognitive load.	- Fixation duration time - Overview and alternating view

Fig. 6. Design recommendations for workflow-oriented desktop VR spaces.

of the overview points and the yellow paths represent the interaction pathways throughout the space, that we identified based on the video analyses. In this mode, users sought a position for each block where all materials related to the block were well-visible and comprehensible. Additionally, an alternating pattern was also evident, similar to displaying parallel windows in 2D. In this case, users held two or more windows in a single view and toggled between active windows with clicks to perform tasks, thus saving time, navigation efforts, and energy.

VI. Discussion

The longer one needs to focus on a particular area, the greater the mental effort required for the user to complete the task or other common activity. In the presented measurement, the mean fixation duration of individuals participating in 2D and 3D measurements was observed. Upon analysis, we found that the mean fixation duration of individuals solving tasks in the 3D VR space was significantly lower than the control group working on the traditional 2D homepage. At the end of each questionnaire, participants were asked to rate the difficulty of the task block on a Likert scale, and at the end of the measurement, a closing questionnaire was administered asking which questionnaire they found most challenging. The questionnaires revealed that although individuals solving VR tasks rated each block slightly more difficult than the control group solving tasks in 2D, there was no significant difference in correctness of the answers and in completion time between the two groups. Furthermore, the mean fixation duration did not support this subjective evaluation; in fact, it indicated the opposite. At the beginning of the measurement, all users were familiar with the use and basic functioning of the 3D software. However, it became clear during the survey and evaluation of the results that due to the layout and design of the space, there was a need for precise navigation between individual blocks in the virtual space, and they were not able to find only a few preference points to see through the whole space, which could negatively affect the user experience. Additionally, observing the users' movements in the space, it is evident that the size and placement of the displays holding digital content in the 3D space also require reconsideration in the future. Although the questionnaires were always centrally located with the necessary content around them, the size and placement varied within the blocks. It would be advisable in the future to restructure these elements so that users are assisted by similar layouts during task completion. For example, videos, PDFs, etc., could be located in the same position in each block and of similar size.

Furthermore, a recurring pattern was that users treated the images containing the titles of blocks as content rather than labels. Therefore, another suggestion for future design is to fully separate the titles/content elements of the content blocks to prevent users from attempting interaction with them and to serve solely as information. Interpreting the results and analyzing user behavior suggests that further optimization and modification of the designed 3D virtual space are necessary to maximize the user experience. To achieve this, we propose the following guidelines.

VII. Designing the Future: Guidelines for VR Development

In virtual environments, user experience hinges on control and presence, crucial elements highlighted by research [32]–[34]. Users' ability to exercise agency in virtual spaces mirrors their actions in the real world, fostering engagement and a sense of security [35], [36]. Active participation and control not only enhance user engagement but also significantly contribute to improved information retention [37], [38]. While tactile feedback in virtual reality may not fully replicate real-world sensations, strategic implementations such as haptic feedback through controllers or gloves aim to enrich the immersive experience [43]. Similarly, auditory cues complement visual elements, aiding users in understanding actions and enhancing their sense of presence [44], [45].

Consistency in design is paramount across various viewing angles and orientations in virtual environments to ensure a seamless user interaction [46], [47].

In previous research within the field of CogInfoCom an cVR, studies [24], [48] have identified key preference points in desktop VR workspaces. In addition, the current study's finding of a significantly shorter fixation duration time in 3D VR underlines the importance of spatial instructions and visual signals in navigating virtual environments. This suggests immersive experiences are more effective than extensive textual guidance [49], [50] in enhancing user commitment. Furthermore, incorporating familiar design elements into these spaces not only increases user comfort but also facilitates task completion [51], [52] [50], [51].

Based on the results of this study on interaction patterns and previous research results of the author regarding the workflow oriented VR spaces, we conclude that effective grouping and clustering of content are essential in virtual workspaces to enhance user experience and effectiveness [24], [48]. Circular arrangements of content groups support holistic overview modes, aiding task monitoring and navigation [24], [53]. Considering the size of virtual spaces is vital to prevent cognitive overload and time loss, with predefined spatial elements facilitating user preferences and customization [53].

Content types, including PDFs, images, presentations, videos, and web content, serve varying roles in virtual environments. Users tend to display PDFs on monitors and videos on projector screens, with display orientation impacting user preference, favoring vertical or slightly inclined displays [53]. Introducing flexible display panels allows users to customize layout, number, size, and relationships of placed content, further enhancing user interaction and customization [53].

Summarizing the main findings above, Figure 6. shows the design recommendations of workflow-oriented VR spaces based on the research results of the current study.

By adhering to these recommendations, VR designers can create spaces that are not only immersive and engaging but also intuitive and efficient for users, leveraging the unique capabilities of VR while addressing its current limitations.

VIII. Conclusion

In conclusion, this study highlights the importance of finding the right balance between providing relevant information and managing potential distractions in 3D VR environments. While the inclusion of detailed 3D animations and well-organized content blocks can enhance immersion, it also introduces complexities that may hinder user experience. Despite users' feelings of increased difficulty, objective eye-tracking data suggest that users navigate 3D VR environments effectively, even in the presence of distractions.

Moving forward, further research is needed to refine strategies for optimizing information presentation in 3D VR environments, with a focus on improving user experience and task performance. Simplifying design principles and prioritizing intuitive navigation can help maximize efficiency and user satisfaction in future 3D VR environments. This study underscores the importance of ongoing research and iterative design processes to fully realize the potential of 3D VR technology in the field of the design of digital workflow-oriented desktop VR spaces.

Acknowledgments

The research presented in this paper was supported by the Hungarian Research Network (HUN-REN), and was carried out within the HUN-REN Cognitive Mapping of Decision Support Systems research group.

References

[1] A. Afrooz, L. Ding, and C. Pettit, "An immersive 3d virtual environment to support collaborative learning and teaching," *Computational Urban Planning and Management for Smart Cities 16*, pp. 267–282, 2019. DOI: 10.1007/978-3-030-19424-6_15

[2] N. Pellas, S. Mystakidis, and A. Christopoulos, "A systematic literature review on the user experience design for game-based interventions via 3d virtual worlds in k-12 education," *Multimodal Technologies and Interaction*, vol. 5, no. 6, p. 28, 2021. DOI: 10.3390/mti5060028

[3] J. M. Evangelista Belo, A. M. Feit, T. Feuchtner, and K. Grønbæk, "Xrgonomics: facilitating the creation of ergonomic 3d interfaces," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–11. DOI: 10.1145/3411764.3445349

[4] Á. B. Csapó, I. Horvath, P. Galambos, and P. Baranyi,"Vr as a medium of communication: from memory palaces to comprehensive memory management," in *2018 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, 2018, pp. 000 389–000 394. DOI: 10.1109/CogInfoCom.2018.8639896

[5] B. Berki, "2d advertising in 3d virtual spaces," *Acta Polytechnica Hungarica*, vol. 15, no. 3, pp. 175–190, 2018. DOI: 10.12700/APH.15.3.2018.3.10

[6] S. Negi and R. Mitra, "Fixation duration and the learning process: An eye tracking study with subtitled videos," *Journal of Eye Movement Research*, vol. 13, no. 6, 2020. DOI: 10.16910/jemr.13.6.1

[7] A. Poole and L. J. Ball, "Eye tracking in hci and usability research," in *Encyclopedia of human computer interaction*. IGI global, 2006, pp. 211–219. DOI: 10.4018/978-1-59140-562-7.ch034

[8] R. N. Meghanathan, C. van Leeuwen, and A. R. Nikolaev, "Fixation duration surpasses pupil size as a measure of memory load in free viewing," *Frontiers in human neuroscience*, vol. 8, p. 1063, 2015. DOI: 10.3389/fnhum.2014.01063

[9] B. Albert and T. Tullis, *Measuring the User Experience: Collecting, Analyzing, and Presenting UX Metrics*. Morgan Kaufmann, 2022. ISBN 0124157920, 9780124157927

[10] S. H. Cha, C. Koo, T. W. Kim, and T. Hong, "Spatial perception of ceiling height and type variation in immersive virtual environments," *Building and Environment*, vol. 163, p. 106285, 2019. DOI: 10.1016/j.buildenv.2019.106285

[11] I. Horváth and B. Berki, "Investigating the operational complexity of digital workflows based on human cognitive aspects," *Electronics*, vol. 12, no. 3, 2023. [Online]. Available: https://www.mdpi.com/2079-9292/12/3/528 DOI: 10.3390/electronics12030528

[12] P. Baranyi and Á. Csapó, "Definition and synergies of cognitive infocommunications," *Acta Polytechnica Hungarica*, vol. 9, no. 1, pp. 67–83, 2012.

[13] P. Baranyi, A. Csapo, and G. Sallai, *Cognitive infocommunications (coginfocom)*. Springer, 2015. DOI: 10.1007/978-3-319-19608-4

[14] I. Horvath, Á. B. Csapó, B. Berki, A. Sudar, and P. Baranyi, "Definition, background and research perspectives behind 'cognitive aspects of virtual reality'(cvr)," *Infocommunications Journal: A Publication of the Scientific Association For Infocommunications (HTE)*, no. SP, pp. 9–14, 2023. DOI: 10.36244/ICJ.2023.SI-IODCR.2

[15] I. Horváth and A. Sudár, "Factors contributing to the enhanced performance of the maxwhere 3d vr platform in the distribution of digital information," *Acta Polytechnica Hungarica*, vol. 15, no. 3, pp. 149–173, 2018. DOI: 10.12700/APH.15.3.2018.3.9

[16] P. Baranyi, P. Galambos, Á. Csapó, and L. Jaloveczki, "Cognitive navigation and manipulation (coginav) method," Feb. 1 2018, uS Patent App. 15/658, 579.

[17] I. K. Boda, E. Tóth, M. Bényei, and I. Csont, "A three-dimensional virtual library model of the ancient library of alexandria," 2015. [Online]. Available: https://api.semanticscholar.org/CorpusID:56215123 DOI: 10.14794/ICAI.9.2014.1.103

[18] I. K. Boda, E. Tóth, I. Csont, and L. T. Nagy, "Toward a knowledge base of literary content focusing on the ancient library of alexandria in the three dimensional space," *2015 6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, pp. 251–258, 2015. [Online]. Available: https://api.semanticscholar.org/CorpusID:25320574 DOI: 10.1109/CogInfoCom.2015.7390600

[19] K. Chmielewska and A. Gilányi, "Mathability and computer aided mathematical education," in *2015 6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, 2015, pp. 473–477. DOI: 10.1109/CogInfoCom.2015.7390639

[20] K. Chmielewska, A. Gilányi, and A. Łukasiewicz, "Mathability and mathematical cognition," in *2016 7th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, 2016, pp. 000245–000 250. DOI: 10.1109/CogInfoCom.2016.7804556

[21] J. B. Kóczy and L. I. Komlosi, "Revisiting literacy: Changing learning paradigms in digital culture," 2019. [Online]. Available: https://api.semanticscholar.org/CorpusID:204873640

[22] C. S. Lányi and Á. Tilinger, "Multimedia and virtual reality in the rehabilitation of autistic children," in *Computers Helping People with Special Needs*, K. Miesenberger, J. Klaus, W. L. Zagler, and D. Burger, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 22–28. DOI: 10.1007/978-3-540-27817-7_4

[23] J. Li, L. Khoo, and S. Tor, "Desktop virtual reality for maintenance training: an object oriented prototype system (v-realism)," *Computers in Industry*, vol. 52, no. 2, pp. 109–125, 2003. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0166361503001039 DOI: 10.1016/S0166-3615(03)00103-9

[24] A. Sudar and A. Csapa, "Interaction patterns of spatial navigation and smartboard use in vr workspaces," in *Accentuated Innovations in Cognitive Info-Communication*. Springer, 2022, pp. 149–166. DOI: 10.1007/978-3-031-10956-0_7

[25] E. D. Ragan, A. Sowndararajan, R. Kopper, and D. A. Bowman, "The effects of higher levels of immersion on procedure memorization performance and implications for educational virtual environments," *PRESENCE: Teleoperators and Virtual Environments*, vol. 19, pp. 527–543, 2010. [Online]. Available: https://api.semanticscholar.org/CorpusID:17361861 DOI: 10.2312/EGVE/JVRC09/121-128

[26] B. Csapó, I. Horváth, P. Galambos, and P. Baranyi, "Vr as a medium of communication: from memory palaces to comprehensive memory management," in *2018 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, 2018, pp. 000 389–000 394. DOI: 10.1109/CogInfoCom.2018.8639896

[27] A. Cannavò and F. Lamberti, "A virtual character posing system based on reconfigurable tangible user interfaces and immersive virtual reality," in *Smart Tools and Applications in Graphics*, 2018. [Online]. Available: https://api.semanticscholar.org/CorpusID:88479188 DOI: 10.2312/STAG.20181297

[28] B. Berki, "Better memory performance for images in maxwhere 3d vr space than in website," in *2018 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, 2018, pp. 000 281–000 284. DOI: 10.1109/CogInfoCom.2018.8639956

[29] ——, "Role of presence, memory and spatial ability in a desktop virtual reality," in Accentuated Innovations in *Cognitive Info-Communication*. Springer International Publishing, sep 2022, pp. 79–97. DOI: 10.1007/978-3-031-10956-0_4

[30] ——, "Desktop vr as a virtual workspace: a cognitive aspect," *Acta Polytechnica Hungarica*, vol. 16, no. 2, pp. 219–231, 2019. DOI: 10.12700/aph.16.2.2019.2.13

[31] I. Horváth, "An analysis of personalized learning opportunities in 3d vr," in *Frontiers of Computer Science*, 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:237587652 DOI: 10.3389/fcomp.2021.673826

[32] S. Weech, S. Kenny, and M. Barnett-Cowan, "Presence and cybersickness in virtual reality are negatively related: a review," *Frontiers in psychology*, vol. 10, p. 158, 2019. DOI: 10.3389/fpsyg.2019.00158

[33] S. Grassini, K. Laumann, and M. Rasmussen Skogstad, "The use of virtual reality alone does not promote training performance (but sense of presence does)," *Frontiers in psychology*, vol. 11, p. 1743, 2020. DOI: 10.3389/fpsyg.2020.01743

[34] N. Cooper, F. Milella, C. Pinto, I. Cant, M. White, and G. Meyer, "The effects of substitute multisensory feedback on task performance and the sense of presence in a virtual reality environment," *PloS one*, vol. 13, no. 2, p. e0191846, 2018. DOI: 10.1371/journal.pone.0191846

[35] A. Tapal, E. Oren, R. Dar, and B. Eitam, "The sense of agency scale: A measure of consciously perceived control over one's mind, body, and the immediate environment," *Frontiers in psychology*, vol. 8, p. 1552, 2017. DOI: 10.3389/fpsyg.2017.01552

[36] M. Velmans, "How could conscious experiences affect brains?" *Journal of Consciousness Studies*, vol. 9, no. 11, pp. 3–29, 2002.

[37] C. Pérez-Sabater, B. Montero-Fleta, M. Pérez-Sabater, B. Rising, and U. De Valencia, "Active learning to improve long-term knowledge retention," in *Proceedings of the xii simposio internacional de comunicación social*, 2011, pp. 75–79.

[38] J. G. Kooloos, E. M. Bergman, M. A. Scheffers, A. N. Schepens-Franke, and M. A. Vorstenbosch, "The effect of passive and active education methods applied in repetition activities on the retention of anatomical knowledge," *Anatomical sciences education*, vol. 13, no. 4, pp. 458–466, 2020. DOI: 10.1002/ase.1924

[39] N. Jayant, J. Johnston, and R. Safranek, "Signal compression based on models of human perception," *Proceedings of the IEEE*, vol. 81, no. 10, pp. 1385–1422, 1993. DOI: 10.1109/5.241504

[40] J. Diemer, G. W. Alpers, H. M. Peperkorn, Y. Shiban, and A. Mühlberger, "The impact of perception and presence on emotional reactions: a review of research in virtual reality," *Frontiers in psychology*, vol. 6, p. 26, 2015. DOI: 10.3389/fpsyg.2015.00026

[41] G. Freeman and D. Maloney, "Body, avatar, and me: The presentation and perception of self in social virtual reality," *Proceedings of the ACM on human-computer interaction*, vol. 4, no. CSCW3, pp. 1–27, 2021. DOI: 10.1145/3432938

[42] F. El Jamiy and R. Marsh, "Survey on depth perception in head mounted displays: distance estimation in virtual reality, augmented reality, and mixed reality," *IET Image Processing*, vol. 13, no. 5, pp. 707–712, 2019. DOI: 10.1049/iet-ipr.2018.5920

[43] H. Benko, C. Holz, M. Sinclair, and E. Ofek, "Normaltouch and texturetouch: High-fidelity 3d haptic shape rendering on handheld virtual reality controllers," in *Proceedings of the 29th annual symposium on user interface software and technology*, 2016, pp. 717–728. DOI: 10.1145/2984511.2984526

[44] M. Hoppe, J. Karolus, F. Dietz, P. W. Woźniak, A. Schmidt, and T.- K. Machulla, "Vrsneaky: Increasing presence in vr through gait-aware auditory feedback," in *Proceedings of the 2019 CHI conference on human factors in computing systems,* 2019, pp. 1–9. DOI: 10.1145/3290605.3300776

[45] A. U. Batmaz and W. Stuerzlinger, "The effect of pitch in auditory error feedback for fitts' tasks in virtual reality training systems," in *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2021, pp. 85–94. DOI: 10.1109/VR50410.2021.00029

[46] S. Bodker, *Through the interface: A human activity approach to user interface design*. CRC Press, 2021. DOI: 10.7146/dpb.v16i224.7586

[47] B. Shneiderman, C. Plaisant, M. Cohen, S. Jacobs, N. Elmqvist, and N. Diakopoulos, *Designing the user interface: strategies for effective human-computer interaction*. Pearson, 2016.

[48] A. Sudár and Á. Csapó, "Interaction patterns of spatial navigation in vr workspaces," in *2019 10th IEEE international conference on cognitive infocommunications (CogInfoCom)*. IEEE, 2019, pp. 615–618. DOI: 10.1109/CogInfoCom47531.2019.9089998

[49] T. Kojić, D. Ali, R. Greinacher, S. Möller, and J.-N. Voigt-Antons, "User experience of reading in virtual reality—finding values for text distance, size and contrast," in *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 2020, pp. 1–6. DOI: 10.1109/QoMEX48832.2020.9123091

[50] P.-L. P. Rau, J. Zheng, Z. Guo, and J. Li, "Speed reading on virtual reality and augmented reality," *Computers & Education*, vol. 125, pp. 240–245, 2018. DOI: 10.1016/j.compedu.2018.06.016

[51] A. M. Cleary, A. S. Brown, B. D. Sawyer, J. S. Nomi, A. C. Ajoku, and A. J. Ryals, "Familiarity from the configuration of objects in 3-dimensional space and its relation to déjà vu: A virtual reality investigation," *Consciousness and cognition*, vol. 21, no. 2, pp. 969–975, 2012. DOI: 10.1016/j.concog.2011.12.010

[52] A. Almutawa and R. Ueoka, "The influence of spatial awareness on vr: investigating the influence of the familiarity and awareness of content of the real space to the vr," in *Proceedings of the 2019 3rd International Conference on Artificial Intelligence and Virtual Reality,* 2019, pp. 26–30. DOI: 10.1145/3348488.3348502

[53] A. Sudár and Á. B. Csapó, "Elicitation of content layout preferences in virtual 3d spaces based on a free layout creation task," *Electronics,* vol. 12, no. 9, p. 2078, 2023. DOI: 10.3390/electronics12092078

Ildikó Horváth obtained her PhD degree in 2017 and in 2023 completed her habilitation in Information Sciences. Her research focuses on the examination of the cognitive aspects of 3-dimensional alternative reality (VR, AR, Mixed reality) systems. The research goal is to determine the directions of development in digital ICT environments that ensure users' work with the most optimal cognitive load, thus positively impacting work efficiency. It also investigates the effects of digital environments on changes in human abilities and the possibilities of personalization. Shas experience in leadership, organization of education and development, and over 10 years of experience in scientific research, publishing, and university education for BSc, MSc, and PhD levels.

Anna Sudár obtained her PhD degree in 2023 in the field of computer science and holds an MSc degree in cognitive science. Her research topic is spatial cognition and the examination of general cognitive abilities in virtual reality. She has more than fifteen scientific publications including a book chapter. In the past 6 years as a scientific community activity, she participated in the organization of several international IEEE conferences and helped the editorial processes of many international Special Journal Issues.

Guidelines for our Authors

Format of the manuscripts

Original manuscripts and final versions of papers should be submitted in IEEE format according to the formatting instructions available on
https://journals.ieeeauthorcenter.ieee.org/
Then click: "IEEE Author Tools for Journals"
- "Article Templates"
- "Templates for Transactions".

Length of the manuscripts

The length of papers in the aforementioned format should be 6-8 journal pages.
Wherever appropriate, include 1-2 figures or tables per journal page.

Paper structure

Papers should follow the standard structure, consisting of *Introduction* (the part of paper numbered by "1"), and *Conclusion* (the last numbered part) and several *Sections* in between.
The Introduction should introduce the topic, tell why the subject of the paper is important, summarize the state of the art with references to existing works and underline the main innovative results of the paper. The Introduction should conclude with outlining the structure of the paper.

Accompanying parts

Papers should be accompanied by an *Abstract* and a few *Index Terms (Keywords)*. For the final version of accepted papers, please send the short cvs and *photos* of the authors as well.

Authors

In the title of the paper, authors are listed in the order given in the submitted manuscript. Their full affiliations and e-mail addresses will be given in a footnote on the first page as shown in the template. No degrees or other titles of the authors are given. Memberships of IEEE, HTE and other professional societies will be indicated so please supply this information. When submitting the manuscript, one of the authors should be indicated as corresponding author providing his/her postal address, fax number and telephone number for eventual correspondence and communication with the Editorial Board.

References

References should be listed at the end of the paper in the IEEE format, see below:

a) Last name of author or authors and first name or initials, or name of organization
b) Title of article in quotation marks
c) Title of periodical in full and set in italics
d) Volume, number, and, if available, part
e) First and last pages of article
f) Date of issue
g) Document Object Identifier (DOI)

[11] Boggs, S.A. and Fujimoto, N., "Techniques and instrumentation for measurement of transients in gas-insulated switchgear," IEEE Transactions on Electrical Installation, vol. ET-19, no. 2, pp.87–92, April 1984. DOI: 10.1109/TEI.1984.298778

Format of a book reference:

[26] Peck, R.B., Hanson, W.E., and Thornburn, T.H., Foundation Engineering, 2nd ed. New York: McGraw-Hill, 1972, pp.230–292.

All references should be referred by the corresponding numbers in the text.

Figures

Figures should be black-and-white, clear, and drawn by the authors. Do not use figures or pictures downloaded from the Internet. Figures and pictures should be submitted also as separate files. Captions are obligatory. Within the text, references should be made by figure numbers, e.g. "see Fig. 2."
When using figures from other printed materials, exact references and note on copyright should be included. Obtaining the copyright is the responsibility of authors.

Contact address

Authors are requested to submit their papers electronically via the following portal address:
https://www.ojs.hte.hu/infocommunications_journal/about/submissions

If you have any question about the journal or the submission process, please do not hesitate to contact us via e-mail:

Editor-in-Chief: Pál Varga – pvarga@tmit.bme.hu

Associate Editor-in-Chief:
Rolland Vida – vida@tmit.bme.hu
László Bacsárdi – bacsardi@hit.bme.hu

Infocommunications Journal

Special Issue

Joint Special Issue on Cognitive Infocommunications and Cognitive Aspects of Virtual Reality

Guest Editors

Péter Baranyi, Atsushi Ito,
Ádám Csapó, Ildikó Horváth,
Tibor Guzsvinecz

Authors

Zoltán Gál, Erzsébet Tóth,
Zakaria Khoudi, Mourad Nachaoui,
Soufiane Lyaqini, Cecilia Sik-Lányi,
Bence Halmosi, Jinat Ara, Judit Szűcs,
Veronika Szűcs, Robert Demeter,
József Katona, Attila Kővári,
Borbála Berki, István Károly Boda,
György Persa, Jinshan Luo,
Yuki Okaniwa, Natsuno Yamazaki,
Yuko Hiramatsu, Madoka Hasegawa,
Péter Földesi, Eszter Sós, Anna Sudár

IEEE ComSoc – IEEE Communications Society

hte – Scientific Association for Infocommunications Hungary – Founded in 1949

IEEE Hungary Section

SCIENTIFIC ASSOCIATION FOR INFOCOMMUNICATIONS

Who we are

Founded in 1949, the Scientific Association for Infocommunications (formerly known as Scientific Society for Telecommunications) is a voluntary and autonomous professional society of engineers and economists, researchers and businessmen, managers and educational, regulatory and other professionals working in the fields of telecommunications, broadcasting, electronics, information and media technologies in Hungary.

Besides its 1000 individual members, the Scientific Association for Infocommunications (in Hungarian: HÍRKÖZLÉSI ÉS INFORMATIKAI TUDOMÁNYOS EGYESÜLET, HTE) has more than 60 corporate members as well. Among them there are large companies and small-and-medium enterprises with industrial, trade, service-providing, research and development activities, as well as educational institutions and research centers.

HTE is a Sister Society of the Institute of Electrical and Electronics Engineers, Inc. (IEEE) and the IEEE Communications Society.

What we do

HTE has a broad range of activities that aim to promote the convergence of information and communication technologies and the deployment of synergic applications and services, to broaden the knowledge and skills of our members, to facilitate the exchange of ideas and experiences, as well as to integrate and harmonize the professional opinions and standpoints derived from various group interests and market dynamics.

To achieve these goals, we…

- contribute to the analysis of technical, economic, and social questions related to our field of competence, and forward the synthesized opinion of our experts to scientific, legislative, industrial and educational organizations and institutions;
- follow the national and international trends and results related to our field of competence, foster the professional and business relations between foreign and Hungarian companies and institutes;
- organize an extensive range of lectures, seminars, debates, conferences, exhibitions, company presentations, and club events in order to transfer and deploy scientific, technical and economic knowledge and skills;
- promote professional secondary and higher education and take active part in the development of professional education, teaching and training;
- establish and maintain relations with other domestic and foreign fellow associations, IEEE sister societies;
- award prizes for outstanding scientific, educational, managerial, commercial and/or societal activities and achievements in the fields of infocommunication.

Contact information

President: **FERENC VÁGUJHELYI** • *elnok@hte.hu*

Secretary-General: **GÁBOR KOLLÁTH** • *kollath.gabor@hte.hu*

Operations Director: **PÉTER NAGY** • *nagy.peter@hte.hu*

Address: H-1051 Budapest, Bajcsy-Zsilinszky str. 12, HUNGARY, Room: 502

Phone: +36 1 353 1027

E-mail: *info@hte.hu*, Web: *www.hte.hu*