IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 29 January 2024, accepted 20 February 2024, date of publication 23 February 2024, date of current version 1 March 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3369587

RESEARCH ARTICLE

Human-Centered Digital Sustainability: Handling Enumerated Lists in Digital Texts

MÁRIA CSERNOCH, TÍMEA NAGY, KEVE NAGY, JÚLIA CSERNOCH, AND CAROLIN HANNUSCH
Faculty of Informatics, University of Debrecen, 4028 Debrecen, Hungary

Corresponding author: Mária Csernoch (csernoch.maria@inf.unideb.hu)

This work was supported in part by the Kdp-2021 Program of the Ministry for Innovation and Technology from the Source of the National Research, Development and Innovation Fund; and in part MTA under Grant 89/2/2022/KP.

ABSTRACT In advance to the present study, the authors introduced a method which makes it possible to calculate the entropy of natural language digital texts, focusing on word-processed texts, presentations, and webpages. This entropy reveals that the more underdeveloped documents are, the more demanding their content-related modification becomes. It was also found that the time and data required to complete a modification task in an erroneous document is several times more than in its correct counterpart. This finding leads to the end-user paradox: the less trained end-users are, the more errors they make, and the modification of their documents requires more resources. To resolve these discrepancies, the present study defines the sustainability rate of natural language digital texts which calculates the losses – the waste of human resources, time, workspace, computers, energy, frustration, working in bees, losing data – generated by negligent text management. Furthermore, we present examples of how manual and enumerated lists behave to modifications in a 213-page long document and conclude from our investigations that while the waste of human and machine resources occurs repeatedly in erroneous documents, the sustainability rate remains low. To prove the necessity of correction, we cleared the sample document, which took approximately 67 hours of two experts of our research group (2 × 67 hours). With this method, we found that the correction of errors can be extremely demanding, but uses resources only once, and further modifications in the now correct document need only the content-required amount of time, activities, entropy, and resources, in accordance with the expectations of the person intended to update the document. To correct documents, we present the Error Recognition Model, which is proved effective and efficient in digital education. All our findings indicate that both education and industry should adapt the presented approach (1) to develop students' and end-users' computational thinking skills, (2) to manage and take advantage of errors, (3) to recognize connections between the structure of the text and the complex word processing tools, (4) to pay attention to digital sustainability – beyond hardware and software development and recycling – with a focus on the human factor. Recently, the Error Recognition Model is a reactive problem-solving approach, whose effectiveness is justified. However, the near future is to run parallel the reactive and proactive uses of this approach, while if we look far into the future, the proactive use to digital born natural language texts should dominate.

INDEX TERMS Documentation, education, human–computer interaction, information entropy, sustainability.

I. INTRODUCTION

A. END-USERS

How can things which are so influential – end-users, end-user activities and products – be ignored in the digital era?

The associate editor coordinating the review of this manuscript and approving it for publication was Liang-Bi Chen.

A huge industry is built upon end-users, but hardly anyone is concerned about how effectively and efficiently they apply both the methodologies and the tools developed for them, and how efficiently resources assigned to these activities are used.

"Despite the widespread use of spreadsheets and the risks and opportunities they pose for organizations, spreadsheets are like dark matter in physics—invisible to IT professionals, corporate managers, and information systems (IS) researchers.'' [1].

Primarily, tech companies – both those producing hardware and those creating software – and many more accompanying industries make profit on end-users, their lack of digital competencies, and low-level computational thinking skills. However, the question is not how much profit is gained by certain groups but how much waste is generated, unnoticed or noticed.

Education is one of the primary concerns of end-user analyses since tech companies have been flooding end-users with their tools and user-friendly slogans but have been unable to produce real user-friendly methods. In general, digital education is a push production system which by definition builds up waste [2], [3], [4], [5], [6]. On the one hand, education gains a lot, since providing courses and accompanying materials makes a huge industry flourish. On the other hand, education is the biggest loser in this situation, since no significant improvement can be detected in end-user activities and production.

In general, end-users have limited vision to tech products [7] and ''with no effective, general education in computational thinking, most people can easily be misled, and they are'' [7]. However, whatever the sources of this low vision and misdirection may be, all participants in the industry – but especially those involved in hardware, software, and education – should accept what has been said by Roter: ''…we cannot see what is coming, a vision based exclusively on current paradigms, competences, products, or technologies can limit the future range of our adaptation too much. Toward that end, a vision should probably focus more on the customers, and broad-scale customer needs, than on ourselves'' [5]. Toyota claims that they want to build better cars for more people [2]. We claim that we want to offer better computer education and skills for more people required for the future of jobs [8], [9], not only for the chosen groups (e.g., IT professionals).

Summary of the section: With the help of this section, we want to sensitize the reader for the importance of digital text management and call attention to its downgrading.

B. DEVELOPMENT OF THE ANALYZED MODEL

To offer more effective and efficient human-centered computer education, we integrate results from our previous studies and new developments presented in this paper.

Previous research proved that modifying erroneous and correct digital texts requires more, respectively, less time. It is found that the time required to carry out a modification depends on the nature of modification: adding data vs. formatting. To demonstrate the impact of modification, three different genres of texts were selected consisting of plain texts, tables, pictures, and mathematical expressions

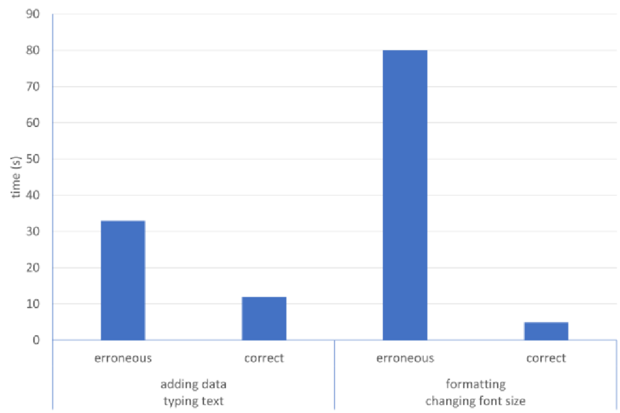

FIGURE 1. Time required in the erroneous and correct documents when adding pieces of text and changing font size were carried out in the medicine document.

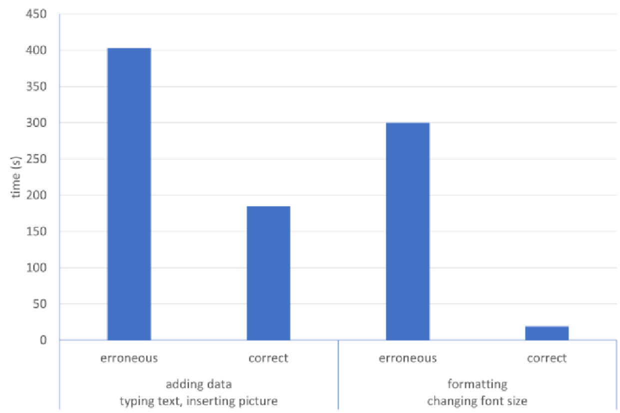

FIGURE 2. Time required in the erroneous and correct documents when adding pieces of text and pictures and changing font size were carried out in the frenchfood document.

(the names of the documents refer to their contents and are used for easier identification):

- medicine: text
- frenchfood: text, table, picture
- equation: mathematical expression

To measure the time required to complete modifications, we need both the original (erroneous) documents and their correct counterparts, which fulfill the requirements of the correctly edited and formatted text (detailed in Section F). This implies that in advance to the recording of the modification, the erroneous documents must be cleared (corrected). The clearing process requires a complete PDCA (Plan-Do-Check-Act) cycle [10] which includes the analysis of the original document, the designing of the clearing process and setting up its algorithm, the testing of the planned movements, and finally the clearing process. The time required for clearing the documents can be estimated by the reading speed where a 100 wpm is accepted in accordance with [11]. In general, we can conclude that the clearing process is time consuming, but it must be carried out because repeated modifications regenerate the losses in erroneous documents.

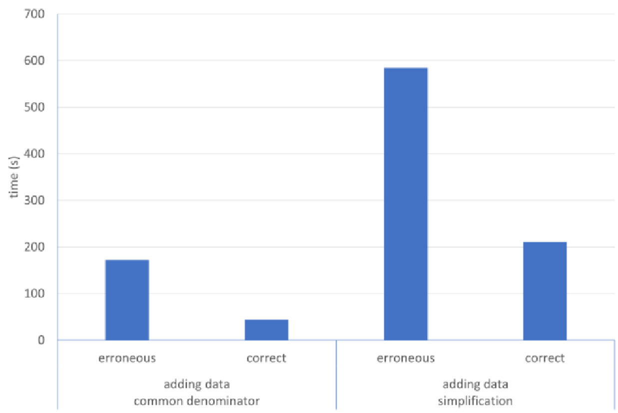

FIGURE 3. Time required in the erroneous and correct documents when adding common denominator and simplification of equations were carried out in the equation document.

Without going into details, Fig. 1–3 demonstrate that in all cases the modifications require more time in the erroneous documents. It is also remarkable but can be derived from the nature of modification, that a simple formatting task (changing font size) is more demanding compared to adding data.

Based on these results, our research group ventured into the calculation of the data required to carry out modifications, detailed in previous publications [12], [13] and the calculation of sustainability rate of modifications, detailed in the present paper.

Summary of the section: The modification of erroneous digital texts is time consuming, and until these errors prevail tremendous resources are squandered. We propose the sustainability rate of digital text to calculate the waste so created. This is the primary concern of the present paper, and the details follow.

C. POSITIONING DIGITAL TEXT MANAGEMENT

Untrained end-users in both internet searching and handling natural language digital texts might not find sources connected to the subject. However, they cannot be blamed for the lack of results since the subject is hardly recognized and studied. If spreadsheets are "like dark matter in physics" [1] then digital texts are darker than one could imagine. Digital text management is something less recognized and researched, downgraded as digital literacy, left out from informatics [14], [15], while "the ability to use the word processor" is identified as the most important skill in computer usage [16]. DigCompEdu [17] goes even further and makes a distinction "…between digital resources and data. Digital resources in this respect comprise any kind of digital content that is immediately understandable to a human user, whereas data need to be analyzed, treated and/or interpreted to be of use for educators." Which implies that the essence of data management is ruled out by DigCompEdu.

Still, one can ask what the difference is between spreadsheets and digital texts. An explanation is that the losses which erroneous spreadsheets can generate have an impact (loss) directly expressed in currency. On the one hand, digital texts do not cause immediate financial losses. On the other hand, to measure the losses that negligent digital text management and digital texts generate, researchers must set up complex objective measuring systems.

To our knowledge, mostly efficiency islands [6] can be found (e.g., for writers: [18], [19], [20], for work and education: [21], [22], [23], [24], [25], [26], [27], [28], [29], [30]), and until now three systems were built on the premises of conceptual problem-solving and its forerunners.

The first was published as early as mid-80s claiming that "Most users are not primarily interested in learning how the system works; they do not want to read through a thick manual or work through an on-line tutorial. Rather, they want to be active; they want to use the system. This orientation, however, can lead to a variety of errors and a lot of time spent recovering from errors. …" [31]. This approach seems to be the forerunner of the Error Recognition Model [32] but disappeared without any traces and/or impacts.

The second system – published around 2000 [30], [37] – still focuses on word processing. Time passed, word processing became more and more popular, but this approach and research results also disappeared unnoticed. The third system, more complex than the previous ones, is the recently published entropy of digital texts which can measure the information content required by modifications regardless of the content, the errors, and the type of document [12], [13]. Hopefully, this system will survive, it will be able to call attention and offer solutions to the problems of handling digital texts. The present paper is based on the entropy of digital texts and its supporting systems (mentioned in the following sections).

It is also remarkable that in 1993 Soloway claimed that programming should be extended to end-user computing [38] and in 2006 Stevens et al. visioned that "…end-users will generally not program in a conventional programming language but will use higher-level means of adaptation that can do the job but are otherwise as simple as possible." [39] However, similar to Ben-Ari's idea [30], [37], these ones also remained unnoticed and/or ignored. One explanation might be found in Grudin's paper reasoning that "As an application matured and use became routine, it got less attention. When email and word processing ceased being discretionary for most of us, CHI researchers moved to the discretionary use frontier of the moment…" [40] which again strengthens the push production approach to computer tools. This discrepancy is further strengthened by the results of the survey of McKenley & Company (questioning 18,000 people in 15 countries) which, on the one hand, found that the respondents' proficiency was lowest in two skill groups in the digital category – 'software use and development' and 'understanding digital systems' – [8]. On the other hand, the

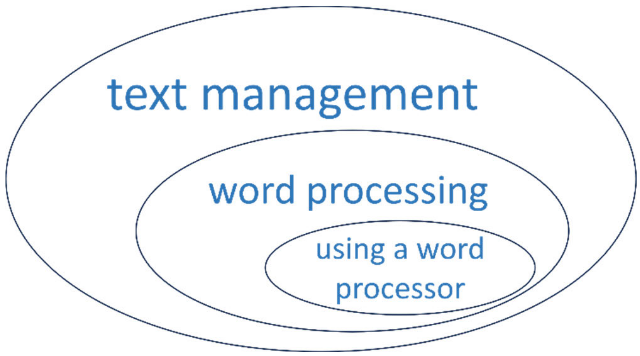

FIGURE 4. The connection between text management activities in word processing environments.

survey also revealed that the seven core skills in 2023 are 'analytical thinking', 'creative thinking', 'resilience, flexibility and agility', 'motivation and self-awareness', 'curiosity and lifelong learning', 'technological literacy', and 'dependability and attention to detail' [9].

In such austerity one might turn to published user-guides, helps, course books, training materials, handouts from teachers, etc. However, the analysis of these materials reveals that only a very limited number of learning-teaching materials consider real-world problem-solving [31], [32], [33], [34], [35], [41], [42], [43], [44], [45]. The quality of most of the resources is questionable for at least two reasons: (1) they are push production materials [46], [47], [48], [49], [50], [51], [52], [53], [54], [55], [56], [57], [58], [59], [60], [61], [62], [63], [64], [65], [66], [67], [68], [69], [70], [71], [72], [73], [74] focusing on the software tools and interfaces and/or (2) prepared by folk teachers [73], [74], [75], [76], [77], [78], [79], [80], [81] – defined by Lister [82] --, who are not necessarily aware and interested of published research results [83], [84] and not rarely unaware of the ineffectiveness of self-studies [31], [85] (e.g., a paper on sustainability [86]).

Furthermore, we must call attention to the difference between using a word processor, word processing, and digital text management. In general, text management is hardly known. Even if it is recognized, the three concepts are used as synonyms. However, this is not the case. The relationship between the three concepts is presented in Fig. 4. The analysis of digital texts reveals that word processing is not enough to create error-free documents, since word processing according to the Encyclopedia of Information Systems is only the act of using a computer to transform written, verbal, or recorded information into typewritten or printed form [87].

The concept of text management primarily focuses on problem-solving, the handling of a content through an iterative process (discussed in Section G). To complete a digital text, one must carry out word processing activities. To do word processing, one must use a word processor. In general, learning a word processor does not cover all the knowledge and skills which are required to create a correct document effectively and efficiently. It usually turns out as bricolage opposed to conceptual models [30], [37]. (According to Ben-Ari, ''bricolage is used to denote a concrete as opposed to an abstract learning style. Turkle and Papert (1990) propose that bricolage be encouraged in computer science education as a way of making the subject accessible to a broader range of students.'' [30])

Nowadays, one can turn to artificial intelligent tools, like ChatGPT [88], in the hope that the huge data inventory behind them might provide useful data and information on erroneous digital texts. In the following, the results of one prompt presented in two languages (English and another, the latter translated to English for the sake of the present paper) are discussed.

For both languages, we entered the same prompt: ''Show me examples for incorrect digital texts.'' Both languages provided scattered, unconnected, inconsistent, unorganized data as answers. In general, the information content of the answers was extremely low. However, a pattern is recognizable in both languages. The English output is about syntax errors – like typos, grammatical errors, spelling, punctuation, capitalization, abbreviation – and two types of semantic errors – incomplete sentences and incorrectly named formatting issues, when there is no Space character between words (for the categories of errors see Section F, Fig. 5, [89]). In the other language the focus is rather on the content, the semantics of the text in the forms: (1) formation of sentences – confusing and overly complex language, long and branching sentences, inaccurate or confusing instructions, incomplete or unsatisfactory content, unnecessary and repetitive information, excessive technical jargon – and (2) tools supporting the understanding of content – few illustrations and visual aids, confusing formatting and style. Furthermore, the error ''unnecessary and repetitive information'' is found in both languages.

These AI supported answers also prove that no valuable data are available on our targeted subject and primarily misconceptions are circulating.

Summary of the section: In this section, we introduce digital text management as an object of research.

D. SCIENTIFIC RESEARCH, FREEDOM OF EXPRESSION, AND RESPECT FOR PRIVATE LIFE

As part of our research, we have a periodically updated set of real-life digital text documents for our use, collected through various channels. One such channel is the unorganized vastness of the internet, offering all kinds of correct and erroneous examples. Unfortunately, most of these documents are published without any accompanying detail regarding who may use them or what one may use them for.

Art. 27 of the Universal Declaration of Human Rights [90] in § 1-2 declares ''Everyone has the right freely to participate in the cultural life of the community, to enjoy the arts and to share in scientific advancement and its benefits.'' When analyzing and correcting the elements of our corpus of digital texts as part of our scientific research, freedom of expression (art. 10 [91]) should be granted. The article specifies "Everyone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers. (...)" [92]. Moreover, the International Covenant on Civil and Political Rights [93] recognizes freedom of expression as a right that can be exercised "either orally, in writing or in print, in the form of art, or through any other media of [the individual's] choice" (art 19, para II). However, when exercising one's freedom of expression, potential violation of other fundamental rights needs to be avoided.

Summary of the section: In this section the legal background is presented, primarily to demonstrate that no rights are violated related to this study.

E. VERY FEW WISHES TO BECOME A BAD EXAMPLE

Our experience shows that apart from a very few devoted pre-service and practicing teachers, people do not like and do not want to be a bad example. Under these peculiar conditions – to keep our results accurate and realistic without throwing dirt on anybody in particular – we apply special care to properly anonymize our samples and otherwise render published examples into something impossible to identify. We may or may not know the exact person or group creating a specific document, but this identity is never relevant for us, nor should it be to our editors or publishers either. This is why we do not mention or disclose such details anywhere, to anybody. What can be relevant for us is only a very generalized subset of the creator's details. For example, what age-range they belong to, what education level they are at, what sort of IT training they possess, what business / industry / education / government field they are working in, or where in a corporate hierarchy they are positioned.

Based on these criteria, our research group set up an introductory fairy tale (Section H) to conceal the real authors and their supervising authority.

Summary of the section: We do not want to point a finger at a determined person/country/company/institution/etc. Our experience shows that the handling of digital texts is independent from all locations, language, age, situation, etc. Therefore, we invented a fairy tale to sensitize the reader to the problem and to present the consequences of ineffective digital text management.

II. ERROR RECOGNITION MODEL (ERM)

"Failures and errors can provide valuable knowledge that can be used in a positive way. The knowledge that is gained from failure and errors can be quite important and may have a big impact on an organization. Failure can augment our experience, knowledge, and resilience. As long as we learn, adapt, and don't repeat the same errors, failure may be an option." [94].

The Error Recognition Model [32] is a complex system to recognize, correct, and eliminate errors in natural language digital texts. It focuses on word-processed texts, presentations, and webpages, but many other documents might be considered with a certain amount of text (e.g., spreadsheets, online quizzes, formative assessment tools, computer games). Another feature of ERM is that it is an enquiry-based learning-teaching approach where both teachers and students can ask questions to guide the processes, and primarily teachers help students to find the right tracks. Finally, it is a "training wheel" system [31] where teachers select the sample texts which should fulfil the following requirements (similar to Carroll's guided exploration cards) [34]: match the students' age, background knowledge, interests (subject integration and Technological Pedagogical Content Knowledge (TPCK) [95] are the keywords),

- developing in complexity at the pace of students (kaizen plays a crucial role) [96], [97], [98], [99],
- varying contents to keep students interested and make them apply earned knowledge in various situations,
- allowing space for repetition and practice to feel safety and familiarity (to call schemata, Fig. 6, Level 3) and to build up schemata (Fig. 6, Level 4).

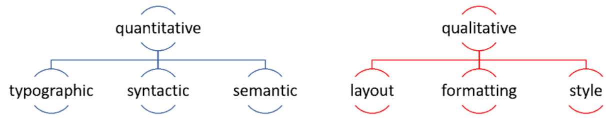

FIGURE 5. The categorization of errors in digital texts.

A. INTRODUCTION TO ERM

The Error Recognition Model (ERM) [32] is based on the following concepts:

- definition of the correctly formatted text,
- categorization of errors in digital texts,
- reading and understanding of both printed and editable andons (visual tools to call attention to errors) [100], [101], [102], [103],
- application of poka-yoke (to prevent the occurrence of error, i.e. a method of designing a fail-safe system) [2], [3], [4], [5], [6].

The definition of the properly formatted text has two criteria:

- The text fulfils the requirements of printed documents (quantitative requirements).
- The content is invariant to modification (qualitative requirements): the document is editable, but the only allowed changes are those matching the original intention of the user.

In accordance with the two criteria of the definition, the errors of the digital texts are categorized in two hypernym and three hyponym categories in each.

In an ERM learning-teaching environment, handling digital texts is introduced with the error recognition process, where the errors of the documents should be recognized, marked, and categorized. The process has two well-distinguishable phases:

TABLE 1. The steps of a concept based problem-solving approach, thinking mode addressed to these steps, and their position in the computer problem-solving typology.

Aim	Type of Thinking	Computer Problem-Solving
understand problem	slow	Level 5
design solution	slow	Level 5
build up schemata	slow	Level 4
call schemata	fast	Level 3
test	fast and slow	all levels

- error recognition in the printed text, including various display formats (e.g., paper, monitor, t-shirt, ads),
- error recognition in the editable format, if available, the andons offered by the software are utilized (e.g., displaying nonprintable characters, gridlines, anchors).

After the errors are recognized and understood, the target condition is to eliminate the errors, to create an error-free document, which is the application of poka-yoke. In this phase, in accordance with the nature of the errors, primarily deleting and unformatting activities are carried out. Actually, this is the first phase when a software tool is used. However, it must be emphasized that only a very limited set is applied.

Finally, when an error-free document is created, the formatting takes place. This is the phase where strictly those tools of the software are selected and applied which are requested to reach the desired, planned format of the text in complete accordance with the content.

It must be emphasized that both the correction and formatting of the digital texts are carried out in a pull system, which implies that not the software interface is taught in full details, but only those tools which are necessary to reach the target condition in an effective and efficient way.

Summary of the section: In this section we introduce the Error Recognition Model (ERM), which is a complex system to handle erroneous digital texts.

B. THE ROLE ERM PLAYS

However, ERM is a reactive response to handle ill-treated digital texts. The ERM process applied to a document is "only" a target condition to reduce waste in the modification process of an already existing document. The goal is to find proactive responses to handle created (new) problems. To fulfil this aim we are required to learn the fundamentals of preparing correct natural language digital texts. This should be a solution-based approach to problem-solving proved to be effective and efficient in either education or industry [7], [10], [104], [105].

These approaches are primarily non-linear, iterative processes which require fast or slow thinking applied appropriately in a situation (Table 1) and can be correctly positioned in the typology of Computer Problem-Solving at the top three levels (Levels 2 and 1 are not real problem-solving approaches but surface management) (Fig. 6).

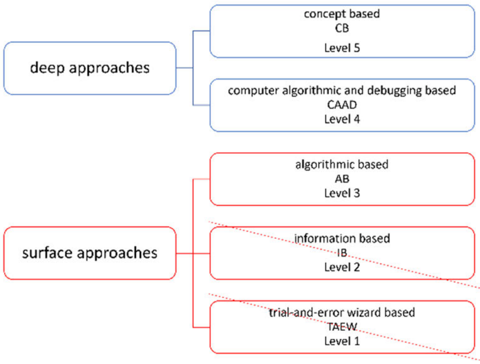

FIGURE 6. The typology of computer problem-solving approaches.

In practice we need a well-defined poka-yoke [2], [3], [4], [5], [6] system which serves the prevention or elimination (yoke) of error or mistake (poka) due to negligence, ignorance, and/or overconfidence. It is intended to prevent the occurrence of error, i.e. a method of designing a fail-safe system. ERM proved as an effective and efficient poka-yoke system in digital text management. Furthermore, due to our knowledge, ERM is the only system built for end-users to handle natural language digital texts in their complexity effectively and efficiently.

Summary of the section: It is proved that ERM offers an effective and efficient approach to handle natural language digital texts.

C. THE KING AND THE DIGITAL REPORT

We would like to emphasize the importance of the current topic by taking the reader on a journey through a little fairy tale. Any resemblance to real persons, companies or countries is purely coincidental. The story is based on a real document available on the internet (see figures in the following sections). However, on the one hand, we do not want to reveal the author of the original document in the present work. On the other hand, creating fictive stories based on real documents and in programming exercises is a kind of gamification, which we practice in class, and students find motivating and funny.

Once upon a time there was a country named Digitalia. Here lived a very old king who was a digital immigrant, according to Prensky [106], [107], but who knew that after his death Digitalia could only survive if they rebuilt their digital strategies. His most talented advisors collected their ideas and created a document of 213 pages. They proudly presented the document to their king. The king was delighted with the achievement of his advisors but suggested some minor modifications.

The king believed that he had not much time left and wanted to finish this report as soon as possible. Therefore, he called for a meeting where all his wise advisors [108] needed to be present and could make the modifications unanimously and efficiently. First, he suggested dividing Chapter 3 into two. The king thought that the modification of the document would take a couple of minutes, after which they could focus on the content. However, chapters were manually numbered, and no matter what the advisors tried, it took hours and hours, and they still could not find a solution with which all of them would have been satisfied. They started to re-number the chapters, but no matter how careful they were, they always missed or duplicated a number somewhere, and then they had to restart the numbering. When this was finally ready, one of them realized that the table of contents had to be renumbered too. Straggling with the table of contents took further hours' work. When finally, they were ready with the table of contents, one of them found that the cross-references also needed to be renumbered. They started browsing the document and wanted to find cross-references. This process again took hours.

The poor old king was extremely fatigued and disappointed, but started to read the restructured report, and came up with further suggestions on the content. He wanted an introductory new Chapter 1, and wanted a new explanatory paragraph inserted at the beginning of Chapter 3. The advisors were really angry because they again had to renumber the whole document, the table of contents, and the cross-references due to the king's request. The king seeing that his wise advisors could not solve the numbering issue, decided to call for his messengers and send them all across the country to find someone who could solve the problem. The king promised half of his kingdom to the person who could provide a solution. Thousands of scholars arrived with their licenses and course books, but none of them was able to find a solution which satisfied the king [109], [110], [111], [112]. He wanted a smart document where changes to the content adjust numbering accordingly.

In a remote village, the news reached a young schoolteacher who, with his students, worked using ERM (Error Recognition Model) [32], [89], [113], [114]. One of their best companions were errors, because they learned that errors are not bad, but something to learn from [2], [3], [5], [99], [115], [116], and [117]. They learned how to recognize and categorize errors [32], [89], how to correct them, and how to handle digital texts efficiently without errors. Their favorite occupation was error hunting [32]. They accepted that errors are pals, and people should not fear errors [5], [99], [115], [116], [117].

The young schoolteacher was convinced that ERM [32] would help the old king. So, he set out on the road. When he reached the castle, no one wanted to let him in, saying that if a problem could not be solved by wise advisors and scholars from all across the country, then how could a schoolteacher do it. The king overheard the commotion in front of the gate and became curious. He found the armed guards fighting a determined young man, who had nothing but a computer to protect himself with. The old king was so disappointed that he let the young schoolteacher in and allowed him to work on the text.

The document was prepared in such a way that changing one thing would imply the automatic change of connected items, e.g., if the numbering #1 – #10 should be changed to #2 – #11. Then inserting a new item before #1 results in the change $1 \mapsto 2, 2 \mapsto 3, \ldots, 10 \mapsto 11$ in one step.

Summary of the section: The fairy tale is based on a real document of a real institution (real situation) which is freely and easily accessible on the internet. We invented this little fairy tale to sensitize the reader to everyday situations and to ensure the anonymity of both the authors and the institution. All matches with real persons and institutions are purely coincidental and will not be verified by us.

D. OVERVIEW OF CONCEPTS

In the following, a table of text management approaches – including word processing and using word processors (Fig. 4) – is presented (Table 2). As the listed sources indicate, the dominant approaches and method are surface approach methods [118] (Fig. 6), which focus on tools, interfaces, primarily in toy situations [35], if there is any situation presented. Others focus on a well-defined group of end-users (e.g., writers, office workers, students) or a specific feature (e.g., non-printing characters). The course book of WordPerfect [41] was among the first teaching-learning materials which offered a deep approach method to text management, but still handling toy situations. We have found the works of Carroll, Ben-Ari, and some open-minded computer scientists who recognized the importance of deep approaches and problem-based learning in end-user activities. The Error Recognition Model, as detailed above handles text management in its complexity, offer a deep learning method with a kaizen approach [2], [3], [4], [5], [6], [96], [97], [98], [99] to match the levels and requirements of students and/or end-users open to handle digital texts effectively and efficiently.

Table 2 uses the following abbreviations:

- problem-based learning (PBL)
- error recognition (ER)
- error categorization (EC)
- non-traditional programming (NTP)
- authentic sources (real situation) (AS)
- definition of correct document (DoCD)

Summary of the section: In this section the major features of teaching-learning materials connected to text management, word processing, and processors are compared.

E. DIGITAL SUSTAINABILITY

For the purposes of the present study, we use the umbrella term of digital sustainability – presented in the paper of Stuermer et al. [122] --, whose aim is to maximize the use of digitalization for the benefit of sustainable development. It implies that knowledge has to be seen as a resource that itself should be sustainable, to preserve its value for society and ensure that it can permanently contribute to the goals of sustainable development.

TABLE 2. Various concepts dealing with text management, word processing, and word processors.

CONCEPTS	TOOLS	LOSSES	PBL	ANDON	ER	EC	NTP	AS	DoCD
surface approaches [17] [46]–[81] [119]–[121]	×								
writers [18]–[20]	×							×	
office workers, students [21]–[30]	×	×						×	
non-printing characters [100]–[103]	×			×	×				
WordPerfect [41]	×		×	×			×		
deep approaches [38] [39] [108]	×		×		×		×	×	
Carroll [31] [33]–[36]	×	×	×		×			×	
Ben-Ari [30] [37]	×	×	×		×			×	
ERM, WDC [12] [13] [32] [44] [45] [89] [113]	×	×	×	×	×	×	×	×	×

However, in general, when digital sustainability is in question [123], [124], it is primarily associated with the construction of faster, lighter, smaller, more user-friendly, more sophisticated hardware or software. As a second thought, the recycling of electrical and electronic equipment and e-waste is considered [125], [126]. In the context of these two major views, where the lifecycle and energy consumption of items with circuitry or electrical components, power or battery supply are dis-cussed, e-waste is one of the fastest growing and most complex waste streams in the world, affecting both human health and the environment, and proliferating the loss of valuable raw materials.

While these are the most obvious aspects, and naturally all are necessary in reaching the goal of an energy-efficient digital sustainability, we also need to point out that this is not enough. The sustainability issue must be attacked from a novel perspective, considering who these devices are built for, and who these devices are used by and who are responsible for creating 'good quality' data: humans [94], [122], [127]. The role of human resources in sustainability is at least as crucial as hardware and software (both programs and data) and their management. According to Tucker, [127] digital sustainability is not necessarily about digital artifacts. "It can also be something much greater but less tangible: sustainable research as contributing critical perspectives and practices that enrich all future work in the humanities."

Our primary concerns – filling a gap in digital sustainability and digital humanities – are to contribute critical perspectives and practices by the analysis of office data management (office computing), the efficiency of office employees and their employers, the source and consequence of wasted human resources, and finally the educational issues connected with these problems. Analyses from previous studies revealed that the low efficiency of office data-handling is almost never recognized and continues without revealing its nature and the negative financial effects caused [12], [30], [32], [37], [89], [113], [114], [128]. The only exception is spreadsheeting, where errors can cause serious financial loss [1], [129], [130]. Errors in digital natural language texts are slyer, and consequently more difficult to recognize, handle, and correct.

Being able to measure the waste of office computing caused by negligent document-management and the lack of concept behind data-handling procedures, a high-mathability [7], [131], [132], [133], [134] problem-solving approach was set up to create a computing environment for supporting efficient data-handling. This method is applied to analyze and modify the selected report.

To position our research and to define the role it plays in digital sustainability, we apply the ten specific basic conditions for digital artifacts, their surrounding ecosystems, and their contribution to sustainable development [122]. In this context, "in a narrow interpretation, a digital ecosystem consists of all hardware devices, program files, and data files that the user needs to process data. In a wider sense, the ecosystem also comprises the social elements which lead to the creation and use of digital artifacts." (Faulkner and Runde 2013 in [122]). Digital artifacts are defined as entities that consist of strings of 0 and 1, which can be interpreted by technical devices, like computers, to provide some meaning. Thus, digital artifacts have become the basic incarnations of knowledge in our times. Consequently, both data files and computer code files can be subsumed under digital artifacts (Kallinikos et al. 2013 in [122]).

From our point of view the findings of Benkler et al. [122] – the quality of digital artifacts is one of the key areas of ongoing research within peer-production ecosystems – is crucial. Applying the ten basic conditions to natural language digital texts, we can find proof that in general they are negative and detrimental to sustainable development (Table 3).

TABLE 3. The ten basic condition of sustainable digital artifacts applied to natural language digital texts [122].

Condition	Natural language digital texts
elaborateness	There are millions of documents. Most of these documents are erroneous, incomplete or of low quality.
transparent structures	Documents are available in editable formats equipped with or without andons, and many only in readable formats (PDF, paper, etc.). In our point of view, e.g., CVs should be requested in editable formats with andons.
semantic data	In general, no semantic data are available even though programs offer styles, layouts, and further tools to properly structure content. If these tools were properly applied, they would serve as taggers.
distributed location	Millions of digital text-based documents are available on the internet, but a lot more are stored in private locations.
open licensing regime	Even the uploaded digital documents are not necessarily available for further publications and research studies (see Sections D and E).
shared tacit knowledge	As detailed, we are in a state where shared tacit knowledge does more harm than good, since primarily misconceptions and push education production systems are circulating.
participatory culture	The authoring and modification culture of the digital text community is not ideal as many of the editors do not have fundamental knowledge but are overconfident. On the other hand, various valuable computing educational theories are unnoticed and/or ignored.
good governance	End-user computing in general, natural digital language text management should be redefined and concept-based learning-teaching methods and their application should be introduced and spread.
diversified funding	Fundings and support are needed to reach education with concept-based methods. First the level of computational thinking skills of teachers should be developed.
contributing to sustainable development	Digital documents have both direct and indirect effects on all three levels of impacts: Life-Cycle, Enabling, and Structural. Considering this aspect also proves that end-user computing plays a crucial role in digital sustainability.

Summary of the section: In this section we introduced the term sustainability for natural language digital texts with a focus on human resources.

III. HYPOTHESES

In connection with the entropy and sustainability rate of modification tasks in natural language digital texts, the following hypotheses were set up.

H1: Modification tasks require more data (higher entropy) in erroneous documents than in their correct counterpart.

H2: The sustainability rate of erroneous documents is lower than that of their correct analogue.

H3: Correcting a document before repeated modification tasks increases the sustainability rate.

IV. METHODS

A. ERRORS AND WASTE

Previous research has revealed that an extremely high number of erroneous text-based documents are in circulation, are multiplied, and serve as samples both in the office world and in education [12], [30], [32], [37], [89], [113], [114]. Creating, editing, and modifying these documents generate serious financial losses when both human and machine resources are considered. It has been found that the application named ANLITA (Atomic Natural Language Input Tracking Application) [135] and a testing method for building up an objective measurement system (MEDT, Measuring the Entropy of Digital Text) [12], [13], [136], [137] serve our purposes in finding the rate of loss in the handling of erroneous documents.

The focuses of our research are the recognition of the errors of digital texts, the measurement of the waste of human and machine resources, the educational and training issues related to these errors, and the integration of our findings and results into workforce optimization. In the present study, we apply the combination of a method (MEDT), the ANLITA tool [135], and a teaching-learning approach (ERM) [32] to shed light on the problems of the numbering used in the selected digital document.

Summary of the section: In this section we introduced ANLITA which is a custom-design software to record activities in text management processes.

B. SAMPLE

For the present study a 213-page-long digital report was selected, which is an MS Word document overloaded with numbered lists. The choice fell upon this document, because its content and its structure are extremely contradictory, and because MS Word is considered the most popular word processor at the time of our research [138]. Due to the length and nature of the errors in this document, the present study selectively focuses on the following issues only:

- numbering of chapters (serving as section headers),
- numbering of paragraphs (serving as content-paragraphs),
- consequences of modifications, including the Table of Contents (TOC) and cross-references.

The document contains other partially enumerated lists, such as footnotes, figures, and tables. However, handling these is beyond the scope of the present paper.

In accordance with ERM [32] in a teaching-learning scenario, the first phase of handling a document is the analysis, followed by correction, and finally by the proper formatting [89], [113]. However, to compare the erroneous and the correct documents, and to calculate the entropy of the tasks applied to both [12], [136], [137], a fourth phase was executed during the research project. At this stage, first level modifications were applied to the texts, where typing and copying took place [29].

Summary of the section: The aim of this section is to introduce the sample document which is used for the present study.

V. DEFINING THE SUSTAINABILITY RATE OF DIGITAL TEXTS

In general, to measure and compare the entropy of the erroneous and the correct documents, the text-handling processes are recorded by ANLITA [135]. Based on the raw data provided by those recordings, we can tell how long a user spent on performing each atomic step. From the number of steps and the time assigned to them, the entropy of each task can be calculated using Shannon's information theory [136], [137] and the entropy of digital text [12]. Once we have calculated the entropy, we can tell how many bits of data (entropy rounded up) (*Eb*) are required (must be put on the communication channel) to carry out that task. The time spent on a task, and the amount of data required to complete it, can provide the sustainability rate of the task, which we will define in Definition 1. With these values we can calculate the amount of waste connected to negligent text handling and management, when both human and computer resources are considered.

Definition 1: Let us denote the time spent on modifying the correct document by tmC and the time spend on modifying the erroneous document by tmE. Then we define the time quotient (denoted by qtm) as the quotient of both (1).

$$qtm = \frac{tmE}{tmC} \tag{1}$$

Further, we denote the entropy of modification in the correct document by emC and the entropy of modification in the erroneous document by emE. Then we define the entropy quotient (de-noted by qem) as the following (2).

$$qem = \frac{2^{emE}}{2^{emC}} = 2^{emE-emC} \tag{2}$$

and finally, we define the sustainability rate (denoted by Sm) as follows (3).

$$Sm = \frac{1}{qtm \cdot qem} \tag{3}$$

Theorem 1: Let us assume that qtm and qem are nondecreasing. Then

$$\lim_{qtm\to\infty} Sm = 0 \text{ and } \lim_{qem\to\infty} Sm = 0$$

Proof: Since qtm and qem are nondecreasing, thus

$$\lim_{qtm\to\infty} qtm = \infty \text{ and } \lim_{qem\to\infty} qem = \infty$$

In a correct document, the sustainability rate of a modification task is 1(Sm), since both qtm and qem are equal to 1.

Therefore, we can conclude that the closer the sustainability rate is to 1, the less erroneous a document is.

We have seen that the entropy of correction tasks is a calculable value [13].

Definition 2: The sustainability rate of a correction task is 0, by definition.

The sustainability rate of a correction task must be 0, because corrections are a waste of time and resources, and can be avoided by creating properly structured and formatted texts [32], [113], [114].

Definition 3: In the case of multiple modifications, we define the cumulative sustainability rate (denoted by S) as the sum of all sustainability rates for each modification divided by n, where n indicates the number of modification tasks (4).

$$S = \frac{\sum_{i=1}^{n} Sm_i}{n} \tag{4}$$

We must mention that in our recordings all correction and modification steps were performed by the members of our research group, who are professional end-users, computer scientists and teachers of informatics. And while the analysis can be repeated by anybody in any digital document, the time and the atomic steps – which are used to calculate the entropy and the sustainability rate – might differ in their case, due to the non-identical background knowledge and skills of the user performing the tasks.

Summary of the section: Here we introduce (define) the sustainability rate of natural language digital texts as a mathematical measure.

VI. AN APPLICATION OF THE SUSTAINABILITY RATE

A. CORRECTION OF NUMBERING IN THE SAMPLE

The numbering of paragraphs is one of the strange characteristics of the selected document. All the paragraphs of the main text are numbered. This is reasonable in chapter headings and sub-headings, but hard to justify or explain in the case of content-paragraphs. The analysis revealed that most of the numbering is manual. And regardless of whether it is manual or automated, only first level numbering is used.

To create proper multilevel numbering, first all the original numbering had to be cleared. We also must note that the original document employed 143 different styles, for which no reasonable explanation was found. All these styles were deleted, and the Normal style was redefined before applying it to the entire document.

Enumerated lists were cleared by applying the Normal style to the entire document, while manual numbering was corrected using the built-in Replace command. Before the series of replacements, a second round of thorough analysis was carried out to reveal the different types of manual numbering in the document. Various patterns were identified, which can be described by the number of leading ordinal symbols (imitation of digits) and the symbols directly following them (Table 4).

In Table 4 t_k denotes the time spent on the event (atomic step) k, I_k denotes the information content of the event k, and E_k denotes the entropy of the event k. Further ^# denotes a digit, ^$ a letter and ^t a tabulator.

The correction of manual numbering was carried out by our expert researchers, who were familiar with the escape and substitute characters of MS Word. Clearing all the manual numbering from the original document took them 188.15 seconds, with 3.1505 entropy (Table 4). All these correction

TABLE 4. The algorithm and the escape sequences of MS Word to replace manual numbering in the document.

step	ESCAPE SEQUENCES	time (t_k)	I_k	E_k
1.	^#^#^#.^t	36.01	2.3862	0.4564
2.	^#^#.^t	11.07	4.0879	0.2404
3.	^#.^t	9.11	4.3691	0.2114
4.	^#^#^#^t	20.00	3.2346	0.3436
5.	^#^#^t	7.02	4.7450	0.1769
6.	^#^t	7.99	4.5583	0.1935
7.	^$^$^$.^t	40.89	2.2028	0.4785
8.	^$^$.^t	12.12	3.9572	0.2548
9.	^$.^t	8.93	4.3978	0.2086
10.	^$. ^t	7.00	4.7492	0.1766
11.	^#)^t	28.11	2.7435	0.4097
	Total	188.15		3.1505

steps and time spent on the analysis are a waste of time and resources, which could have been avoided by a properly designed, structured, and formatted text [13], [89], [113].

The analysis also revealed that both the Table of Contents (TOC) and cross-references were inserted manually. Consequently, each of them had to be located and renumbered when a change in the numbering of the chapters was introduced. To create a properly structured and formatted document, these errors were corrected. The details of the correction of the entire document are beyond the scope of the present paper. However, to calculate the sustainability rate of modification tasks, this correct document was used in comparison to the original (erroneous) document.

Summary of the section: The clearing process of the original erroneous document is outlined. Based on the estimated reading speed for learning and understanding [11] and the recording of the process, we calculated the time. To correct the document took two experts from our research group and not less than 67 hours per person.

B. MODIFICATIONS OF NUMBERED LISTS

To demonstrate how fragile the original document was, and how predictable the correct one became, we performed the same modifications in both documents.

- A new Chapter 1 was inserted in front of the original Chapter 1. With this task we could demonstrate and measure how such a modification affects the numbering of chapters, the TOC, and the cross-references to chapters.
- The first content-paragraph was copied and pasted in front of the original Paragraph 1. With this copying we could prove the immediate need for manual renumbering and the efficiency of automated numbering.

The modification tasks and the calculation of the entropy and sustainability rates were executed in the following order.

- For each modification task, an algorithm was set up.
- Based on the algorithm, the time (t_k) spent on the atomic steps was identified from the data logged by ANLITA [13], [135].
- Using the atomic steps, the probability, the information content (I_k), and the entropy of each step (E_k) were calculated. The sum of the *Ek* values provides the entropy (*E*), and the rounded-up entropy gives us the number of bits (*Eb*) which must be put on the communication channel to solve that problem. The primary sources of this data are education (training) and the user interface of the word processor (GUI).
- Finally, the sustainability rate was calculated from the comparison of the modification tasks between the original and the correct document.

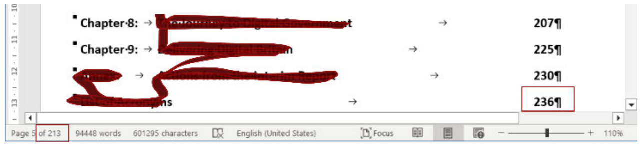

FIGURE 7. Manual numbering in the TOC of the original document.

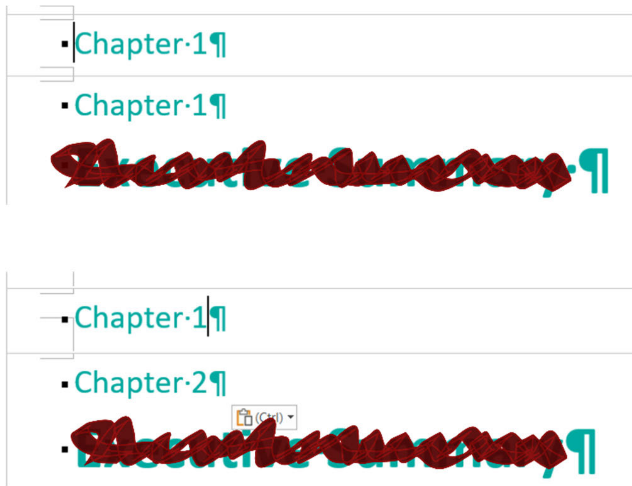

FIGURE 8. Chapter 1 copied in the original document with manual numbering (upper) and in the correct document with automated numbering (lower). The cursor is placed at the frontmost position.

1) CHANGING THE NUMBERING OF THE CHAPTERS

The analysis of the original document revealed that the TOC and the cross-references were constructed manually, the TOC contains non-existing page numbers, and the document is unable to automatically adjust to changes in the numbering of the chapters. The manual TOC is presented in Fig. 7. The screenshot indicates that the page numbers of chapters were not updated, which led to incorrect and non-existing numbers.

On the other hand, if chapters are enumerated and styled automatically, and the TOC and the cross-references are created based on the existing automated numbering and styles, then the word processor is able to track changes and modify the document in accordance. In the samples, the cursor is placed at the first possible position in the line, to demonstrate that in the original document the word "chapter" and its accompanying number are inserted as content (Fig. 8 upper sample), while in the correct document a formatting was applied (Fig. 8 lower sample).

TABLE 5. Inserting a new Chapter 1 into the correct document and its expenditure of time.

ALGORITHM	time (t_k)	I_k	E_k
position cursor	3.96	4.0172	0.2481
insert new paragraph with Normal style	1.05	5.9323	0.0971
selection	2.09	4.9392	0.1610
copy formatting	5.87	3.4493	0.3158
update	20.18	1.6678	0.5249
update TOC	30.97	1.0499	0.5071
Total	64.12		1.8540

In the following table (Table 5), we summarize the time which is needed to insert a new chapter into a correct document. We needed 16 times more time in the original document than in the correct one to insert a new Chapter 1 section heading. A task which can be executed in around 1 minute, requires more than a quarter of an hour in an erroneous document, without any assurance that all the required updates are completed. In the erroneous document there is no guarantee that the quality of the document reaches a high standard.

The algorithms of the modification of Chapter 1 were set up for both the original (erroneous) and the correct documents. In the correct document 6 atomic steps were recognizable, the time spent on these steps was 64.12 s, and the entropy was 1.8540 (Table 5). In the original (erroneous) document, where the update had to be performed manually, to execute the same modification 63 steps and 1048 s were required with an entropy of 4.0874. We can conclude that to complete this modification 2 and 5 bits of data (*Eb*) must be put on the communication channel in the correct and the original document, respectively (which proves H1).

To calculate the sustainability rate of this modification (inserting a new chapter heading) in the original (erroneous) document, we used the recorded time and the calculated entropy. The method and formulae introduced in Section 0 gave us a 0.0078 sustainability rate, which is extremely close to 0.

2) INSERTING A NUMBERED PARAGRAPH

In the next phase, a new numbered paragraph was inserted into the text, just in front of Paragraph 1. Fig. 9 shows the cursor at the first possible position, in the original and the correct document, before and after copying the paragraph.

In the correct document, copying Paragraph 1 required 4 atomic steps which were completed in 6 s with an entropy of 1.9265 ($Eb = 2$). These steps are presented in Table 6. Algorithm of inserting a new numbered content-paragraph into the correct document. along with the time (t_k), the information content (I_k) and the entropy (E_k). To perform the same task in the erroneous text, we needed 150 atomic steps, taking 714 s with an entropy of 7.1244 ($Eb = 8$, which also proves H1).

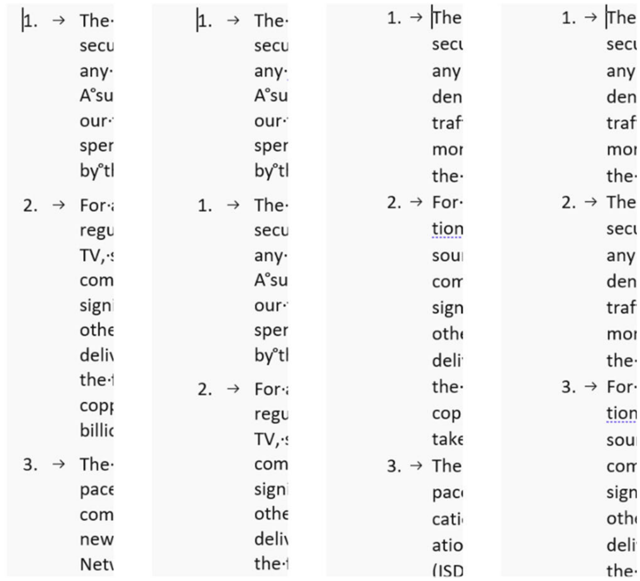

FIGURE 9. The original and the modified manual numbering (l1 and l2), and the original and the modified automated numbering (r2 and r1). The cursor is placed at the frontmost position.

TABLE 6. Algorithm of inserting a new numbered content-paragraph into the correct document.

ALGORITHM	time (t_k)	I_k	E_k
select	1.88	1.6694	0.5248
copy to Clipboard	2.05	1.5445	0.5295
position cursor	1.01	2.5658	0.4334
paste	1.04	2.5236	0.4389
Total	5.98		1.9265

In the next phase of the analyses, we repeated the same alteration ten times in the erroneous, the corrected, and the correct document, recording the process, and measuring the time required to complete the tasks. The result is presented in Fig. 10. It shows that if we modify an erroneous document repeatedly, it can be mapped with a linear approximation.

However, if the errors are corrected, the time of the correction is used once, and after that only the time of modification is added on top.

In the erroneous document, the sustainability rate of this simple task (inserting a numbered content-paragraph) is 0.0001. This result demonstrates that handling manual numbering is a waste of time, human, and machine resources. Furthermore, manual numbering carries the risk of left-out or duplicate numbers, which further lowers the quality of the document.

Calculating and mapping the sustainability rate of the repeated modifications in the three documents (Fig. 11) shows that once the document is correct, only those steps are executed that the user originally planned [113]. There is no waste, because we do not need to execute collateral steps (proves H2 and H3). With this method we also found proof

FIGURE 10. The time required to insert a numbered paragraph in the original, corrected, and correct documents.

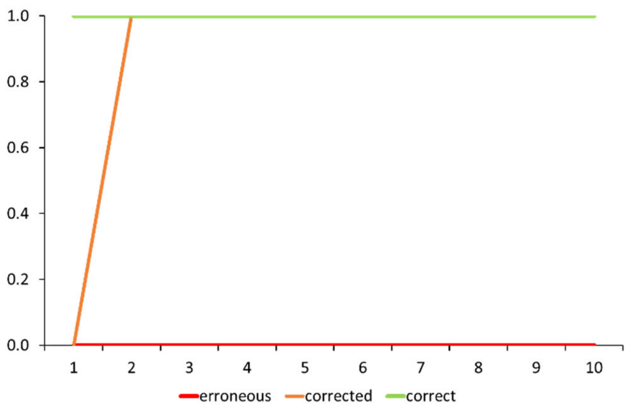

FIGURE 11. The sustainability rate of inserting a numbered paragraph into the original, corrected, and correct documents.

for the statement of Hatamleh & Tilesh "As long as we learn, adapt, and don't repeat the same errors, failure may be an option." [94].

Summary of the section: In this section, we present two real situations to modify enumerated list. Based on these real situations, we calculated three sustainability rates to repeated modifications: in the (1) erroneous, (2) corrected, and (3) correct documents. We can conclude that to act proactively we should avoid errors, to act reactively we must correct documents. Both methods are to eliminate waste and/or to reduce the losses of human and machine resources while handling natural language digital texts.

VII. SUMMARY

A. AND THEY LIVED HAPPILY EVER AFTER

The king was surprised at how the schoolteacher worked. First, he analyzed the text thoroughly [5], [96], [97], [98], [99], [115], [116], and when he recognized the errors of the document, corrected them, and then set up different sets of automated numbering. The method worked, and the king was delighted to find a solution to their problem. He offered half of his kingdom, but the schoolteacher did not accept it, because he wanted to go back to his school and his students. However, in exchange for the half kingdom, he asked the king to find resources to train his advisors, the scholars, and the people of his kingdom, to learn how to do quality work in text editing. He also asked for a research center where the development in human-centered sustainability could be studied.

The king, seeing the inefficiency of his advisors' text-handling activities compared to the method introduced by the schoolteacher, recognized that not seeing the structure of a text, the essence of a complex program, and not connecting the concepts of informatics to handling natural language texts with a word processor leads to a waste of human and machine resources. He also realized that to make changes, education had to be rethought, and methods should be introduced which could increase quality and decrease waste. He also realized that to make changes, small steps should be taken first [5], [96], [97], [98], [99], so he prohibited the use of manual chapter numbering in documents in his country. He recommended asking questions (asking themselves or others, [10], [96], [97], [98], [104], [105]) before one wanted to imitate an enumerated/itemized list by typing numbers/characters. He ordered his subjects to start removing erroneous documents from social media to reduce their spread.

B. HOW TO APPLY ERM?

Many might be surprised by the ERM approach [32], which involves analyzing the text thoroughly first, and when the errors of the document are recognized, correcting them, and then setting up different sets of automated solutions, in the present case, numbering. The approach is well-known in industry [2], [3], [4], [5], [99], [139], and the effectiveness of ERM is proved, and it works in education [32].

Thinking forward, we should find resources to train our students, advisors, clerks, and office workers, to learn how to do quality work in text editing, together with a research center, where the development in human-centered sustainability could be studied. It must be recognized that not seeing the essence of complex programs and not finding the concepts of informatics in handling natural language texts with a word processor lead to a waste of human and machine resources. To make changes, education must be rethought, and methods should be introduced which can increase quality and decrease waste. To achieve our goals, small steps should be taken first [2], [3], [5], [96], [97], [99] and develop slowly. This is a time-consuming process which cannot be rushed.

VIII. DISCUSSION

The primary aims of the present paper are to set up the theoretical background to calculate the sustainability rate of digital natural language texts from the entropy of the modification of tasks, to present examples of how theory can be turned into practice, and to indicate how education can support human-centered digital sustainability.

We have found that the sustainability rate of a task can be derived from the duration, the information content, and the entropy of the atomic steps of the task. This value is in the [0, 1] range: 0 indicating collateral corrections, and 1 the planned modifications of a correct document. In an erroneous document this range is (0, 1). For the present study the numbered lists of a report were analyzed, corrected, properly formatted, and modified. After producing a corrected version of this document, modifications were performed both in the original (erroneous) and the correct document. While the time spent on the analysis of the report is approximate, all the other activities are measured, including the development of macros for repeated actions. All these digital steps of the process were recorded with the logging application ANLITA [135], dedicated to tracking end-user input actions. To test the partial sustainability of the document, two modifications on what-appeared-to-be numbered lists were performed: First, a new numbered chapter was inserted as a section heading. Here the table of contents and cross-references were also checked and modified. Then, a numbered content-paragraph was duplicated, where consecutive paragraphs were also renumbered.

In the sample erroneous document, we found the sustainability rates to be close to 0 for both modification tasks, meaning that the waste of human and machine resources is extremely high, because of end-users' weak computational thinking skills and ignorance of quality digital work. Many believe that the less developed a document is, the easier it is to handle, modify or manage it. This misconception originates from confusing the underdevelopment of user-skills with that of digital products [109], [110]. However, we found proof that the more properly structured and built texts are, the less demanding they are to maintain. This paradox suggests that those end-users create erroneous documents whose computational thinking skills are low, even though handling such content demands more expertise than is needed when handling correct ones. This is a vicious circle that we can only escape by making end-users' computational thinking skills more developed. The results make it obvious that the amount of data required to complete a task does not necessarily imply the quality of the outcome, and low-quality data requires more data. It is also proved that correcting the errors first is highly profitable before applying repeated modifications. These findings are in complete accordance with productivity theories in industry [5], [99].

IX. CONCLUSION AND FURTHER RESEARCH

The target conditions of the present paper are to set up an objective measuring system to natural language digital texts, call attention to the human role in digital sustainability, and offer solutions to reduce the waste of human and machine resources through education. We find this crucial in the development process [89], [140] as has been proved by industry [5], [99]. Our further target condition and/or hope is that we will be able to reach an audience sensitive to digital sustainability and eliminating waste generated by negligent end-user activities.

To accomplish this goal, first the computational thinking skills of teachers – regardless of their subjects – should be developed [140]. Once they are able to recognize the importance of information content and the sustainability of digital texts, they can decide whether a text is properly edited, they can teach the method to their students, and then can expect them to handle digital texts correctly. In this process, cognitive load [104], [141], [142], [143], TPCK (Technological Pedagogical Content Knowledge) [95], and subject integration [12], [144] play crucial roles.

Based on these results, more studies are required to find out how demanding various texts and modification tasks are, but it is already obvious that erroneous text management and the lack of fundamental computational thinking skills use up our resources senselessly. Further research questions are how the sustainability rate can be converted into quality education, how the actual waste can be calculated in currency, and how education and waste can be balanced to make human-centered digital office-work sustainable.

The selected sample document is not unique, several similar – both in length and content – are available on the internet, shared by authors and re-shared by readers who are unaware of the errors which these documents carry. The country Digitalia is invented to disguise the author of the analyzed document, however millions of erroneous documents can be found in our close environment, in our private collections, and on the internet. This includes that sources to carry out sustainability analyses are available in immense quantity. Before claiming that you do not have any erroneous document in your folders, please keep in mind that computer-related self-assessment is rather misleading, the Dunning-Kruger effect takes its toll [26], [109], [146], [147]. The error categories presented in Section F can help you to recognize the errors in digital texts (word-processed texts, presentations, webpages, etc.) [89]. The severity of the errors can be measured by the system – calculating entropy and sustainability rate – detailed in the present paper.

With the presented method, an objective tool is offered to make comparative analysis and to measure the losses generated in both education and industry (in the real world). We can calculate how much time is needed to fulfil a modification task, how many computers how much workspace and energy are needed and wasted due to negligent text handling.

Hopefully, this approach will not end up unnoticed as its predecessors, and based on these results we can call attention to the role and importance of human-centered digital sustainability. Furthermore, we hope to see course books which leave behind the tool-centered push production systems with toy situations, and switch to human-centered, concept-based problem-solving. The Error Recognition Model is such an approach, which has been proved to be effective and efficient in our education system and provided fundamentals to the present research.

Furthermore, we should keep in mind that it is not the software to be blamed for the errors – see Grove's famous misconceptions which lead to banishing text and data management from schools [145] --, but the methods with which we learn, teach, and use these tools. As previous but unnoticed research studies and the present findings proved, conceptualization plays a crucial role in both the learning and the application processes, consequently these approaches should be adapted first in schools, trainings, and in learning-teaching materials [34], [35], [82], [83], [84], [140], [148], and then in industry.

In accordance with the results gained from the definitions and the methods for the calculation of the entropy and the sustainability rate of natural language text, we can analyze both digital documents and human behavior and acts in digital surroundings. We further plan to test real situations in both educational and industrial settings to call attention to the problem and find space for publishing real situation results and the approach in full detail.

ACKNOWLEDGMENT

Supported by the KDP-2021 (Cooperative Doctoral Program, in Hungarian: Kooperatív Doktori Program) Program of the Ministry for Innovation and Technology from the Source of the National Research, Development and Innovation Fund. Carolin Hannusch was supported by the Hungarian Academy of Science's grant for the support of researchers raising children, 89/2/2022/KP.

REFERENCES

[1] R. R. Panko, "The cognitive science of spreadsheet errors: Why thinking is bad," in *Proc. 46th Hawaii Int. Conf. Syst. Sci.*, Jan. 2013, pp. 4013–4022, doi: 10.1109/HICSS.2013.513.

[2] T. Ohno, *Toyota Production System: Beyond Large-Scale Production*. New York, NY, USA: Productivity Press, 1988.

[3] J. F. Krafcik, "Triumph of the lean production system," *Sloan Manage. Rev.*, vol. 30, no. 1, pp. 41–52, 1988.

[4] J. P. Womack and D. T. Jones, *Lean Thinking: Banish Waste and Create Wealth in Your Corporation*. New York, NY, USA: Simon & Schuster, 2003.

[5] M. Rother, *Toyota Kata: Managing People for Improvement, Adaptiveness, and Superior Results*. New York, NY, USA: McGraw-Hill, 2010.

[6] N. Modig and P. Ahlström, *This is Lean. Resolving the Efficiency Paradox*. Stockholm, Sweden: Rheologica Publishing, 2018.

[7] C. Wolfram, *The Math(s) FIX: An Education Blueprint for the AI Age*. Champaign, IL, USA: Wolfram Media, 2020.

[8] M. Dondi, J. Klier, F. Panier, and J. Schubert. *Defining the Skills Citizens Will Need in the Future World of Work*. Future-citizen skills McKinsey. Accessed: Jan. 3, 2024. [Online]. Available: https://www.mckinsey.com/~/media/mckinsey/industries/public%20and%20social%20sector/our%20insights/defining%20the%20skills%20citizens%20will%20need%20in%20the%20future%20world%20of%20work/defining-the-skills-citizens-will-need-in-the-future-of-work-final.pdf

[9] *Future of Jobs Report 2023–World Economic Forum*. Accessed: Jan. 3, 2024. [Online]. Available: https://www3.weforum.org/docs/WEF_Future_of_Jobs_2023.pdf

[10] A. Smalley, *Four Types of Problems: From Reactive Troubleshooting to Creative Innovation*. Cambridge, MA, USA: Lean Enterprise Institute, Inc., 2018.

[11] MTCT. *Speed Reading*. MindTools. Accessed: Jan. 28, 2024. [Online]. Available: https://www.mindtools.com/aokg6bn/speed-reading

[12] M. Csernoch, J. Máth, and T. Nagy, "The interpretation of graphical information in word processing," *Entropy*, vol. 24, no. 10, p. 1492, Oct. 2022, doi: 10.3390/e24101492.

[13] M. Csernoch, K. Nagy, and T. Nagy, "The entropy of digital texts—The mathematical background of correctness," *Entropy*, vol. 25, no. 2, p. 302, Feb. 2023, doi: 10.3390/e25020302.

[14] J. Vahrenhold, E. Nardelli, C. Pereira, G. Berry, M. E. Caspersen, J. Gal-Ezer, A. McGettrick, and M. Westermeier, *Informatics Education in Europe: Are We All in the Same Boat? Report by the Committee on European Computing Education (CECE)*. New York, NY, USA: Association for Computing Machinery, 2017.

[15] W. Gander. (Nov. 21, 2012). *Informatics in Schools? Urgently Needed*. ECSS, Barcelona. Accessed: Jan. 3, 2024. [Online]. Available: https://people.inf.ethz.ch/gander/talks/GanderECSS2012.pdf

[16] W.-W. Jiang, W. Chen, and Y.-C. Chen, "Important computer competencies for the nursing profession," *J. Nursing Res.*, vol. 12, no. 3, pp. 213–226, 2004, doi: 10.1097/01.jnr.0000387505.98877.6d.

[17] C. Redecker, *European Framework for the Digital Competence of Educators. DigCompEdu*, Y. Punie, Ed. Luxembourg: Publications Office of the European Union, 2017.

[18] E. Johnson. *A Little Learning About Word Processing*. Learning About Word Processing. Accessed: Jan. 21, 2023. [Online]. Available: https://www.uv.es/ fores/programa/johnson_wordprocessing2.html

[19] M. G. Kirschenbaum, *Track Changes: A Literary History of Word Processing*. Cambridge, MA, USA: The Belknap Press of Harvard Univ. Press, 2016.

[20] R. T. Kellogg and S. Mueller, "Cognitive tools and thinking performance: The case of word processors and writing," in *Proc. Annu. Meeting Psychonomic Soc.*, 1989, pp. 1–30. [Online]. Available: https://files.eric.ed.gov/fulltext/ED311455.pdf.

[21] S. Gibbs, A. McKinnon, and G. Steel. *Are Workplace End-User Computing Skills at a Desirable Level? A New Zealand Perspective*, AIS Electronic Library (Aisel) AMCIS 2014 Proceedings: Are Workplace end-User Computing Skills at a Desirable Level? A New Zealand Perspective. Accessed: Jan. 22, 2023. [Online]. Available: https://aisel.aisnet.org/amcis2014/EndUserIS/GeneralPresentations/5/

[22] S. F. Gibbs and T. J. McLennan, "Has the computing competence of first year university students increased during the last decade?" in *Proc. 25th Annu. ASCILITE Conf.*, Nov. 2008, pp. 632–640.

[23] S. Gibbs, G. Steel, and A. Kuiper, "Expectations of competency: The mismatch between employers' and Graduates' views of end-user computing skills requirements in the workplace," *J. Inf. Technol. Educ., Res.*, vol. 10, pp. 371–382, Jan. 2011, doi: 10.28945/1531.

[24] S. F. Gibbs, "The above-average effect in an end-user computing context," M.S. thesis, Dept. Tourism, Sport Soc., Lincoln Univ., Christchurch, New Zealand, 2016.

[25] S. Gibbs, "Computer self-efficacy—Is there a gender gap in tertiary level introductory computing classes?" *J. Appl. Comput. Inf. Technol.*, vol. 17, no. 1, pp. 1–6, 2013.

[26] S. Gibbs, K. Moore, G. Steel, and A. McKinnon, "The dunning-Kruger effect in a workplace computing setting," *Comput. Hum. Behav.*, vol. 72, pp. 589–595, Jul. 2017, doi: 10.1016/j.chb.2016.12.084.

[27] S. Gibbs, G. Steel, and A. McKinnon, "A content validity approach to creating an end-user computer skill assessment tool," *J. Appl. Comput. Inf. Technol.*, vol. 19, no. 1, pp. 1–8, 2015.

[28] A. Repenning and A. Ioannidou, "What makes end-user development tick? 13 design guidelines," in *Human-Computer Interaction Series*. Dordrecht, The Netherlands: Springer, 2006, pp. 51–85, doi: 10.1007/1-4020-5386-X_4.

[29] K. Hankiewicz and M. Butlewski, *Efficiency in Performing Basic Tasks Using Word Processing Programs by the Elderly as a Measure of the Ergonomic Quality of Software* (Lecture Notes in Computer Science). Cham, Switzerland: Springer, 2014, pp. 481–488, doi: 10.1007/978-3-319-07233-3_44.

[30] M. Ben-Ari, "Bricolage forever!" in *Proc. 11th Annual Workshop Comput.-Based Learn. Unit*. Leeds, U.K.: Univ. of Leeds, Jan. 1999. [Online]. Available: http://www.ppig.org/papers/11th-benari.pdf

[31] R. Catrambone and J. M. Carroll, "Learning a word processing system with training wheels and guided exploration," *ACM SIGCHI Bull.*, vol. 17, no. SI, pp. 169–174, May 1986, doi: 10.1145/30851.275625.

[32] S. Katalin, C. Gábor, M. Csernoch, and B. Aradi, "Error recognition model: High-mathability end-user text management," *Acta Polytechnica Hungarica*, vol. 19, no. 1, pp. 151–170, 2022, doi: 10.12700/aph.19.1.2022.1.10.

[33] J. M. Carroll, R. L. Mack, C. H. Lewis, N. L. Grischkowsky, and S. R. Robertson, "Exploring exploring a word processor," *Hum.–Comput. Interact.*, vol. 1, no. 3, pp. 283–307, Sep. 1985, doi: 10.1207/s15327051hci0103_3.

[34] J. M. Carroll and M. B. Rosson, *Paradox Act. User*. San Jose, CA, USA: IBM Thomas J. Watson Research Division, 1986.

[35] J. M. Carroll, *Interfacing Thought: Cognitive Aspects of Human-Computer Interaction*. Cambridge, MA, USA: MIT Press, 1987.

[36] J. M. Carroll, "Evaluation, description and invention: Paradigms for human-computer interaction," *Hum. Comput. Interact.*, vol. 29, pp. 1–22, Aug. 1988, doi: 10.21236/ada204617.

[37] M. Ben-Ari and T. Yeshno, "Conceptual models of software artifacts?" *Interacting With Comput.*, vol. 18, no. 6, pp. 1336–1350, Dec. 2006, doi: 10.1016/j.intcom.2006.03.005.

[38] E. Soloway, "Should we teach students to program?" *Commun. ACM*, vol. 36, no. 10, pp. 21–24, Oct. 1993, doi: 10.1145/163430.164061.

[39] G. Stevens, G. Quaisser, and M. Klann, "Breaking it up: An industrial case study of component-based tailorable software design," in *End User Development* (Human-Computer Interaction Series), Nov. 2006, pp. 269–294, doi: 10.1007/1-4020-5386-X_13.

[40] J. Grudin, *A Moving Target-The Evolution of Human-Computer Interaction* (Human-Computer Interaction Handbook), J. Jacko, Ed. Boston, MA, USA: Taylor & Francis, 2012.

[41] *WordPerfect Workbook: For IBM Personal Computers and PC Networks, Version 5.1*, WordPerfect Corp., Orem, UT, USA, 1989.

[42] S. Kinn and T. Siann, "Introduction to word processing," in *Computers and Clinical Audit*. Boston, MA, USA: Springer, 1998, pp. 91–99, doi: 10.1007/978-1-4899-6639-1_8.

[43] A. L. Kamouri, J. Kamouri, and K. H. Smith, "Training by exploration: Facilitating the transfer of procedural knowledge through analogical reasoning," *Int. J. Man-Mach. Stud.*, vol. 24, no. 2, pp. 171–192, Feb. 1986, doi: 10.1016/s0020-7373(86)80047-5.

[44] M. Csernoch, "From webtables to datatables," in *Proc. EuSpRIG Conf. Spreadsheet Risk Manage., EuSpRIG Eur. Spreadsheet Risks Interes*, 2019, pp. 1–22.

[45] K. Sebestyén, G. Csapó, and M. Csernoch, "The effectiveness of the webtable-datatable conversion approach," *Annales Mathematicae et Informaticae*, vol. 56, pp. 109–121, Jan. 2023, doi: 10.33039/ami.2022.12.010.

[46] I. Tsimperidis and G. Chrysomalidis, "Timeless and multi-platform teaching of word processors," *J. Modern Educ. Rev.*, vol. 11, pp. 1287–1293, Sep. 2021.

[47] I. M. Joos, D. M. Wolf, and R. Nelson, "Introduction to word processing," in *Introduction to computers for Healthcare Professionals*. Burlington, MA, USA: Jones & Bartlett Learning, 2021.

[48] J. Lambert, *Microsoft Word 2016*. Redmond, WA, USA: Redmond: Microsoft, 2015.

[49] A. Philo and M. Angstadt, *Microsoft Word 2016 Step-By-Step Guide*. Norristown, PA, USA: MC-NPL Computer Lab, 2020.

[50] L. Bucki, *Microsoft Word 2013 Bible*. Indianapolis, IN, USA: Wiley, 2013.

[51] J. Walkenbach, *Excel 2016 Bible*. Indianapolis, IN, USA: Wiley, 2015.

[52] F. Wempen, *Microsoft PowerPoint 2013 Bible*. Indianapolis, IN, USA: Wiley, 2013.

[53] D. Gookin, *Microsoft Word for Dummies*. Hoboken, NJ, USA: Wiley, 2019.

[54] E. Smyrnova-Trybulska, *Effective Development of Teachers' Skills in the Area of ICT and e-Learning: Monograph*. Katowice, Poland: Studio Noa, 2017.

[55] L. Bucki, J. Walkenbach, F. Wempen, D. Kusleika, and M. Alexander, "Creating documents with word 2013," in *Microsoft Office 2013 Bible*. Indianapolis, IN, USA: Wiley, 2013, pp. 101–398.

[56] A. Lènárd, G. Sarbó, M. T. Éder, and N. Turzó-Sovák, "Digital culture 3," in *Hungarian: Digitális Kultúra 3*. Budapest, Hungary: Oktatási Hivatal, Budapest, 2022.

[57] A. Lènárd, N. Turzó-Sovák, M. T. Éder, and G. Sarbó, "Digital culture 4," in *Hungarian: Digitális Kultúra Tankönyv 4*. Budapest, Hungary: Oktatási Hivatal, Budapest, 2023.

[58] A. Lènárd, A. Abonyi-Tóth, N. Turzó-Sovák, and P. Varga, "Digital culture 5," in *Hungarian: Digitális Kultúra 5*. Budapest, Hungary: Oktatási Hivatal, Budapest, 2020.

[59] A. Abonyi-Tóth, C. Farkas, N. Turzó-Sovák, and P. Varga, "Digital culture 6," in *Hungarian: Digitális Kultúra 6*. Budapest, Hungary: Oktatási Hivatal, Budapest, 2021.

[60] A. Abonyi-Tóth, C. Farkas, and P. Varga, "Digital culture 7," in *Hungarian: Digitális Kultúra 7*. Budapest, Hungary: Oktatási Hivatal, Budapest, 2022.

[61] A. Abonyi-Tóth, C. Farkas, and P. Varga, "Digital culture 8," in *Hungarian: Digitális Kultúra 8*. Budapest, Hungary: Oktatási Hivatal, Budapest, 2023.

[62] P. Varga, "Digital culture 9," in *Hungarian: Digitális Kultúra 9*. Budapest, Hungary: Oktatási Hivatal, Budapest, 2020.

[63] A. Abonyi-Tóth, C. Farkas, Z. Reményi, and K. J. Horváth, "Digital Cculture 10," in *Hungarian: Digitális Kultúra 10*. Budapest, Hungary: Oktatási Hivatal, 2021.

[64] A. Abonyi-Tóth, "Digital culture 11," in *Hungarian: Digitális Kultúra 11*. Budapest, Hungary: Oktatási Hivatal, 2022.

[65] C. D. Baican and M. E. Coriteac, "Informatics and ICT 5," in *Romanian: Informatica Si TIC-Clasa 5-Manual*. Bucureşti, Romania: Sigma, 2018.

[66] D. Popa, "Informatics and ICT 6," in *Romanian: Informatică Şi TIC: Manual Pentru Clasa a VI-a*. Bucureşti, Romania: Editura Didactică Si Pedagogică, 2020.

[67] V. Giurgiulescu and V. B. Giurgiulescu, "Informatics and ICT 6," in *Romanian: Informatica Si TIC-Clasa 6-Manual*. Bucureşti, Romania: Editura Didactică Si Pedagogică, 2018.

[68] L. Ciocaru, S. Penea, O. Rusu, and C. E. Mitrache, "Informatics and ICT 7," in *Romanian: Informatica Si TIC-Clasa 7-Manual*. Bucureşti, Romania: Litera Educational, 2019.

[69] A. Florea and S.-E. Săcuiu, "Informatics and ICT 7," in *Romanian: Informatica Si TIC-Clasa 7-Manual*. Bucureşti, Romania: Editura Didactică Si Pedagogică, 2019.

[70] L. Ciocaru, S. Penea, O. Rusu, and C.-E. Mitrache, "Informatics and ICT 8," in *Romanian: Informatica Si TIC-Clasa 8-Manual*. Bucureşti, Romania: Editura Litera, 2020.

[71] A. Florea and S.-E. Săcuiu, "Informatics and ICT 8," in *Romanian: Informatica Si TIC-Clasa 8-Manual*. Bucureşti, Romania: Editura Didactică Si Pedagogică, 2021.

[72] R. K. Barbhuiya, Ed., *Informatics Practices. Textbook for Class XI*. New Delhi, India: National Council of Educational Research and Training, 2021.

[73] (2023). *NCERT, ICT Textbook-Class IX. Creating Textual Communication, Chapter 2. NCERT Books*. Accessed: Jan. 3, 2024. [Online]. Available: https://ncert.nic.in/textbook/pdf/iict102.pdf

[74] (2023). *NCERT, Word Processing Tool 3. NCERT Books*. Accessed: Jan. 3, 2024. [Online]. Available: https://ncert.nic.in/textbook/pdf/kect103.pdf

[75] Gareth. (Jul. 13, 2017). *Qualification Component and Record of Learner Achievement. Word Processing Software Skills E3 (L/616/1318)*. Accessed: Jan. 3, 2024. [Online]. Available: https://www.aim-group.org.uk/clientfiles/files/units/learner_achievements/Word%20Processing%20Software%20Skills%20E3%20CV2%20AIM%20Awards%20Component%20V1-1.docx

[76] *Building Institutional Capacity From the Heart of Africa*, document BEST PRACTICE AWARDS 2016 ECDL/ICDL, ECDL Foundation, 2016, pp. 31–32.

[77] *Word processing*. Word Processing Unit 2: Manipulating Text With Word Lesson 4: Basic Text Formatting. Accessed: Dec. 29, 2023. [Online]. Available: http://www.webmediaworkshop.com/Classes/cs100/mod.wp/unit.2/lesson.4/textformat.html

[78] *Word Processing*. Word Processing Assignment for Unit 2. Accessed: Dec. 29, 2023. [Online]. Available: http://www.webmediaworkshop.com/Classes/cs100/mod.wp/unit.2/assign2.html

[79] *Word Processing*. Word Processing Unit 2: Manipulating Text With Word Lesson 4: Basic Text Formatting. Accessed: Dec. 29, 2023. [Online]. Available: http://www.webmediaworkshop.com/Classes/cs100/mod.wp/unit.2/lesson.4/paraformat.html

[80] R. McNally and S. O'Neill, "Word processing uses and features," in *For Students of Monaghan Institute*. Monaghan, Ireland, Dec. 2023. [Online]. Available: https://miitnotes.weebly.com/uploads/9/8/4/1/9841023/lo1_word_processing_uses_and_features.pdf

[81] E. Goodwin, "Word Processing basics," in *Bus Technology Application*, Dec. 2023. Accessed: Jan. 3, 2024. [Online]. Available: https://slideplayer.com/slide/6126983/#google_vignette

[82] R. Lister, "After the gold rush: Toward sustainable scholarship in computing," in *Proc. 10th Conf. Australas. Comput. Educ.*, vol. 78, 2008, pp. 3–17.

[83] L. Malmi, J. Sheard, P. Kinnunen, Simon, and J. Sinclair, "Computing education theories: What are they and how are they used?" in *Proc. ACM Conf. Int. Comput. Educ. Res.*, 2019, pp. 187–197, doi: 10.1145/3291279.3339409.

[84] L. Malmi, J. Sheard, P. Kinnunen, Simon, and J. Sinclair, "Development and use of domain-specific learning theories, models, and instruments in computing education," *ACM Trans. Comput. Educ.*, vol. 23, no. 1, pp. 1–48, Mar. 2023, doi: 10.1145/3530221.

[85] *Spreadsheet Competency Framework: A Structure for Classifying Spreadsheet Ability in Finance Professionals*, ICAEW, London, U.K., 2016.

[86] M. R. Green, L. M. Walters, T. Walters, and L. Wang, "Not just another research paper: Understanding global sustainability through digital documentary," *Social Stud.*, vol. 106, no. 1, pp. 37–46, Jan. 2015, doi: 10.1080/00377996.2014.964390.

[87] A. Prestage. *Word Processing*. (Encyclopedia of Information Systems). Amsterdam, The Netherlands: Academic, 2003. [Online]. Available: https://www.sciencedirect.com/topics/computer-science/word-processing

[88] *ChatGPT*. ChatGPT. Accessed: Dec. 31, 2023. [Online]. Available: https://openai.com/chatgpt

[89] M. Csernoch, "Thinking fast and slow in computer problem solving," *J. Softw. Eng. Appl.*, vol. 10, no. 1, pp. 11–40, 2017, doi: 10.4236/jsea.2017.101002.

[90] *Universal Declaration of Human Rights*, United Nations, New York, NY, USA, 2017.

[91] *European Court of Human Rights: European Convention on Human Rights*, Council of Europe, Strasbourg, France, 1950. [Online]. Available: https://www.echr.coe.int/documents/convention_eng.pdf

[92] (2022). *European Court of Human Rights: Guide on Article 10 of the European Convention on Human Rights. Freedom of expression*. Council of Europe. Accessed: Aug. 31, 2022. [Online]. Available: https://www.diva-portal.org/smash/get/diva2:1655124/FULLTEXT01.pdf

[93] *General Assembly Resolution 2200A (XXI): International Covenant on Civil and Political Rights*, United Nations, 1976. Adopted and Opened for Signature, Ratification, and Accession by General Assembly resolution 2200A (XXI) of 16 December 1966, Entry Into Force 23, New York, NY, USA, Mar. 1976.

[94] O. Hatamleh and G. Tilesch, *Between Brains; Taking Back Our AI Future*. Redwood City, CA, USA: GTPublishDrive, 2020.

[95] C. Angeli and N. Valanides, *Technological Pedagogical Content Knowledge—Exploring, Developing, and Assessing TPCK*. Cham, Switzerland: Springer, 2015.

[96] R. Maurer, *The Spirit of Kaizen: Creating Lasting Excellence One Small Step at a Time*. New York, NY, USA: McGraw-Hill, 2013.

[97] R. Maurer, *One Small Step Can Change Your Life: The Kaizen Way*. New York, NY, USA: Workman Pub, 2014.

[98] T. Narusawa and J. Shook, *Kaizen Express: Fundamentals for Your Lean Journey*. Cambridge, MA, USA: Lean Enterprise Institute, 2009.

[99] J. K. Liker, *The Toyota Way: 14 Management Principles From the World's Greatest Manufacturer*. New York, NY, USA: McGraw-Hill, 2004.

[100] S. R. Dodlapati, P. Lakkaraju, N. Tulluru, and Z. Zeng, *Non Printable & Special Characters: Problems and How to Overcome Them*. Accessed: Jan. 22, 2023. [Online]. Available: https://www.lexjansen.com/nesug/nesug10/ff/ff04.pdf

[101] C. K. Kenyon. *Show/Hide Non-Printing Formatting Characters*. Accessed: Jan. 21, 2023. [Online]. Available: https://www.addbalance.com/word/nonprinting.htm

[102] S. Barnhill. *Word's Nonprinting Formatting Marks*. Nonprinting Formatting Marks. Accessed: Jan. 24, 2023. [Online]. Available: http://wordfaqs.ssbarnhill.com/NonprintChars.htm

[103] E. Curts. *How to Show Non-Printing Characters in Google Docs*. LaptrinhX. Accessed: Jan. 21, 2023. [Online]. Available: https://laptrinhx.com/how-to-show-non-printing-characters-in-google-docs-989631243/

[104] G. Polya, *How to Solve It*. Princeton, NJ, USA: Princeton Univ. Press, 1945.

[105] T. Brown, *Change by Design, Revised and Updated: How Design Thinking Transforms Organizations and Inspires Innovation*. New York, NY, USA: HarperCollins Publishers, 2019.

[106] M. Prensky, "Digital Natives, digital immigrants," *On the Horizon*, vol. 9, no. 5, pp. 1–6, 2001, doi: 10.1108/10748120110424816.

[107] M. Prensky, "Digital natives, digital immigrants part 2: Do they really think differently?" *Horizon*, vol. 9, no. 6, pp. 1–6, Nov. 2001, doi: 10.1108/10748120110424843.

[108] J. M. Wing, "Computational thinking," *Commun. ACM*, vol. 49, no. 3, pp. 33–35, Mar. 2006, doi: 10.1145/1118178.1118215.

[109] J. Kruger and D. Dunning, "Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments.," *J. Personality Social Psychol.*, vol. 77, no. 6, pp. 1121–1134, 1999, doi: 10.1037/0022-3514.77.6.1121.

[110] S. Staub and R. Kaynak, "Is an unskilled really unaware of it?" *Proc. Social Behav. Sci.*, vol. 150, pp. 899–907, Sep. 2014, doi: 10.1016/j.sbspro.2014.09.099.

[111] M. Csernoch, P. Biró, J. Máth, and K. Abari, "Testing algorithmic skills in traditional and non-traditional programming environments," *Informat. Educ.*, vol. 14, no. 2, pp. 175–197, Oct. 2015, doi: 10.15388/infedu.2015.11.

[112] T. Nagy, M. Csernoch, and P. Biró, "The comparison of students' self-assessment, gender, and programming-oriented spreadsheet skills," *Educ. Sci.*, vol. 11, no. 10, p. 590, Sep. 2021, doi: 10.3390/educsci11100590.

[113] M. Csernoch, "Teaching word processing—The theory behind," *Teaching Math. Comput. Sci.*, vol. 7, no. 1, pp. 119–137, 2009.

[114] M. Csernoch, "Teaching word processing—The practice," *Teaching Math. Comput. Sci.*, vol. 8, no. 2, pp. 247–262, 2010, doi: 10.5485/tmcs.2010.0252.

[115] B. M. McLaren, D. Adams, K. Durkin, G. Goguadze, R. E. Mayer, B. Rittle-Johnson, S. Sosnovsky, S. Isotani, and M. van Velsen, *To Err is Human, to Explain and Correct is Divine: A Study of Interactive Erroneous Examples With Middle School Math Students* (Lecture Notes in Computer Science). Berlin, Germany: Springer, 2012, pp. 222–235, doi: 10.1007/978-3-642-33263-0_18.

[116] J. Metcalfe, "Learning from errors," *Annu. Rev. Psychol.*, vol. 68, no. 1, pp. 465–489, Jan. 2017, doi: 10.1146/annurev-psych-010416-044022.

[117] D. C. Webb, "Troubleshooting assessment: An authentic problem solving activity for it education," *Proc. Social Behav. Sci.*, vol. 9, pp. 903–907, Jan. 2010.

[118] D. H. J. M. Dolmans, S. M. M. Loyens, H. Marcq, and D. Gijbels, "Deep and surface learning in problem-based learning: A review of the literature," *Adv. Health Sci. Educ.*, vol. 21, no. 5, pp. 1087–1112, Nov. 2015, doi: 10.1007/s10459-015-9645-6.

[119] *Exam Simulation ICDL Module 3: Word Processing*. Test ICDL ECDL & IT Certifications. Accessed: Jan. 27, 2024. [Online]. Available: https://www.testicdl.com/module-3-word-processing.html

[120] *ECDL / ICDL Word Processing Syllabus Version 5.0*. Testicdl. Accessed: Jan. 27, 2024. [Online]. Available: https://www.testicdl.com/images/syllabus-ecdl/ECDL_ICDLWordProcessing1.pdf

[121] *Previous Maturation Exam Periods. Tasks and Revision-Assessment by Subject*. Hungarian: Korábbi érettségi idöszakok Feladatai és Javítási-értékelési útmutatói Vizsgatárgyanként. Accessed: Jan. 27, 2024. [Online]. Available: https://dari.oktatas.hu/erettsegi.utmutato.index

[122] M. Stuermer, G. Abu-Tayeh, and T. Myrach, "Digital Sustainability: Basic conditions for sustainable digital artifacts and their ecosystems," *Sustainability Sci.*, vol. 12, no. 2, pp. 247–262, 2016, doi: 10.1007/s11625-016-0412-2.

[123] P. Messerli and E. Murniningtyas, Eds., *The Future is Now: Science for Achieving Sustainable Development: Global Sustainable Development Report 2019*. New York, NY, USA: UN, Dept. of Economic and Social Affairs, 2019.

[124] L. Jensen, Ed., *Sustainable Development Goals Report 2022*. New York, NY, USA: United Nations, 2022. [Online]. Available: https://unstats.un.org/sdgs/report/2022/The-Sustainable-Development-Goals-Report-2022.pdf

[125] (2022). *IBM Cloud Education*. [Online]. Available: https://www.ibm.com/cloud/blog/green-computing

[126] V. Forti, C. P. Baldé, R. Kuehr, and G. Bel. (2020). *The Global E-Waste Monitor 2020. Quantities, Flows, and the Circular Economy Potential*. UNU/UNITAR and ITU. Accessed: Jan. 23, 2023. [Online]. Available: https://ewastemonitor.info/wp-content/uploads/2020/11/GEM_2020_def_july1_low.pdf

[127] J. Tucker, "Facing the challenge of digital sustainability as humanities researchers," *J. Brit. Acad.*, vol. 10, pp. 93–120, 2022, doi: 10.5871/jba/010.093.

[128] M. Csernoch and E. Dani, "Do you speak, write and think in informatics?" *Acta Polytechnica Hungarica*, vol. 19, no. 1, pp. 113–131, 2022, doi: 10.12700/aph.19.1.2022.1.8.

[129] *European Spreadsheet Risk Interest Group*, EuSpRIG1 (n.d.). Accessed: Jan. 3, 2024. [Online]. Available: https://eusprig.org/

[130] *Horror Stories*. EuSpRIG2 (n.d.). European Spreadsheet Risk Interest Group. Accessed: Jan. 3, 2024. [Online]. Available: https://eusprig.org/research-info/horror-stories/

[131] P. Baranyi and A. Gilányi, "Mathability: Emulating and enhancing human mathematical capabilities," in *Proc. IEEE 4th Int. Conf. Cognit. Infocommunications (CogInfoCom)*, Dec. 2013, pp. 555–558.

[132] P. Baranyi, Á. Csapó, and G. Sallai, *Cognitive Infocommunications (CogInfoCom)*. Cham, Switzerland: Springer, 2015.

[133] P. Biró and M. Csernoch, ‘‘The mathability of computer problem solving approaches,’’ in *Proc. 6th IEEE Int. Conf. Cogn. Infocommunications (CogInfoCom)*, Oct. 2015, pp. 111–114.

[134] C. Wolfram. (May 22, 2015). *Evidence: Let’s Promote Not Stifle Innovation in Education*. Conrad Wolfram. [Online]. Available: https://www.conradwolfram.com/writings/2015/5/21/role-of-evidence-in-education-innovation

[135] K. Nagy and M. Csernoch, ‘‘Pre-testing erroneous text-based documents: Logging end-user activities,’’ *Frontiers Educ.*, vol. 7, Mar. 2023, Art. no. 958635, doi: 10.3389/feduc.2022.958635.

[136] C. E. Shannon, ‘‘A mathematical theory of communication,’’ *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 623–656, Jul. 1948.

[137] C. E. Shannon, ‘‘A mathematical theory of communication,’’ *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul. 1948.

[138] A. Walker. (Dec. 7, 2022). *23 Best Word Processing Software (2023)*. Guru99. Accessed: Jan. 3, 2024. [Online]. Available: https://www.guru99.com/best-free-word-processor.html

[139] A. Altshuler, M. Anderson, D. Jones, D. Roos, and J. Womack, *The Future of the Automobile: The Report of MIT’s International Automobile Program*. Cambridge, MA, USA: MIT Press, 1984.

[140] J. Hattie, *Visible Learning for Teachers: Maximizing Impact on Learning*. Evanston, IL, USA: Routledge, 2012.

[141] J. Sweller, P. Ayres, and S. Kalyuga, *Cognitive Load Theory*. New York, NY, USA: Springer, 2011.

[142] D. Kahneman, *Thinking, Fast and Slow*. New York, NY, USA: Farrar, Straus and Giroux, 2011.

[143] J. Chen, D. Morris, and N. Mansour, ‘‘Science teachers’ beliefs: Perceptions of efficacy and the nature of scientific knowledge and knowing,’’ in *International Handbook of Research on Teachers’ Beliefs*. New York, NY, USA, 2014.

[144] M. Caspersen. (2022). *Informatics Reference Framework for School*. [Online]. Available: https://www.informaticsforall.org/wp-content/uploads/2022/03/Informatics-Reference-Framework-for-School-release-February-2022.pdf

[145] M. Gove. *Digital Literacy Campaign Michael Gove’s Speech in Full*. Guardian. Accessed: Jan. 21, 2023. [Online]. Available: http://www.theguardian.com/education/2012/jan/11/digital-literacy-michael-gove-speech

[146] *Swiss ECDL Study: Poor Basic Digital Skills and Overestimation Used as Marketing Tool*, document BEST PRACTICE AWARDS 2016 ECDL/ICDL, ECDL Foundation, 2016, pp. 66–67.

[147] *Perception & reality of Digital Skills*. ICDL Europe. Accessed: Dec. 30, 2023. [Online]. Available: https://www.icdleurope.org/policy-and-publications/perception-reality-measuring-digital-skills-gaps-in-europe-india-and-singapore/

[148] P. A. Kirschner, J. Sweller, and R. E. Clark, ‘‘Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching,’’ *Educ. Psychologist*, vol. 41, no. 2, pp. 75–86, Jun. 2006, doi: 10.1207/s15326985ep4102_1.

MÁRIA CSERNOCH received the Teacher degree in mathematics, descriptive geometry, informatics and English, the B.Sc. degree in software engineering, the Ph.D. degree in mathematics and computer sciences, and the Dr. (Habil.) degree in applied linguistics from the University of Debrecen, Hungary. She is currently an Associate Professor with the Faculty of Informatics, University of Debrecen. Her research interests include didactics of informatics, specialized in developing computational thinking skills, knowledge-transfer, subject integration, digital sustainability, and lean computing education.

TÍMEA NAGY received the Teacher degree in informatics and mathematics, in 2020. She is currently pursuing the Ph.D. degree with the University of Debrecen, Hungary. She is also a Project Coordinator with Hungarian company specializing in the development of e-learning course materials. Her research interests include didactics of informatics, specialized in developing computational thinking skills, knowledge-transfer, subject integration, and digital sustainability.

KEVE NAGY received the B.Sc. (Technical Teacher) degree from the Polytechnic of Dunaújváros (now University of Dunaújváros) and the M.Sc. (IT Teacher) degree from the University of Veszprém (now University of Pannonia). He is currently pursuing the Ph.D. degree in applied IT with the University of Debrecen. He is also an actively practicing Applied IT Engineer and a continuously Researching Teacher/Trainer. He was educated in a Protestant Secondary School, became a Certified Telecommunications Technician. He tries to balance his life experience between being a computing engineer and teaching.

JÚLIA CSERNOCH received the French-English Translator degree from the University of Debrecen, Hungary, and the Law degree from the University of Neuchatel, Switzerland. She is currently pursuing the Ph.D. degree with the University of Debrecen. Her research interests include contextualization of informatics, subject integration, copyright issues in research, digital sustainability, fundamental human rights, and lean computing.

CAROLIN HANNUSCH received the Ph.D. degree in mathematics from the University of Debrecen, Hungary, in 2018. Currently, she is an Assistant Professor with the Faculty of Informatics, University of Debrecen. Her research interests include coding theory, IT security, and computer science education.