https://doi.org/10.36315/2024v2end001
Education and New Developments 2024

THE VALUES AND LEGAL ISSUES OF AUTHENTIC DATA SOURCES IN COMPUTER EDUCATION AND RESEARCH

Júlia Csernoch, & Mária Csernoch
Faculty of Informatics, University of Debrecen (Hungary)

Abstract

Data management, especially text management is a dubious issue in computing education (CE). Some claim that it is only digital literacy. Others, on the contrary, emphasize that text management is part of CE. Our research reveals that not data management, but the approaches meant to handle the subject are responsible for its displacement, where the issue of meaningless formatting and tools are in the focus. To handle data management with standard software tools, we must give up the widely accepted decontextualized approaches, instead, authentic digital sources must be introduced in classes. To teach and research text management, find and reveal errors, and guide student show to avoid them, readymade but editable documents must be collected with real content. The internet, without question, is the largest source of authentic documents, and teachers must give students access to these documents for analysis, treatment and/or interpretation. However, teaching effective, efficient, and sustainable text management with authentic sources, legal issues arise (e.g., copyright, protection of personal data, freedom of expression). For this reason, documents from various internet sources must be considered as lawfully accessible (three-step test, art. 5 (5) InfoSoc Directive). The present paper details the purpose of using authentic data in teaching data management, analysing various documents closely connected to the subject, and provides examples where law allows, while others block their use.

Keywords: *Authentic document, lawful access, legal issues, copyright, text management.*

1. Introduction

For teachers and researchers, it is imperative that information is transferred to the future generation. The present paper describes the background of an educational pull production system (Csernoch, 2017; Sebestyén et al., 2022) that is not only effective by considering all aspects of TPCK (Technological Pedagogical Content Knowledge) (Angeli & Valanides, 2015) but efficient by eliminating end-user generated soft-waste (Csernoch et al., 2022, 2023; Nagy & Csernoch, 2023). The method uses the technique of collecting existing (authentic) data and learning through data analysis and errors, not only of our own but of others as well (McLaren et al., 2012; Metcalfe, 2016; Tulis et al., 2016). Documents are countless and accessible through the internet, and we are convinced that this accessibility can and should be used in education. However, teachers and researchers must know how to select and use the available sources in both classroom and dissemination.

In our paper, we present that data and erroneous documents can be used in education and even in a very productive and important way. The question remains whether teaching (including scientific publication) through available data, including erroneous documents that are accessible online legally authorized to be analyzed, corrected etc. in a classroom or not. The aim of the present paper is to answer this question, which might seem easy at first, but in reality, it is not evident.

2. Sources of authentic documents

The internet, without question, is the largest source of authentic documents, and teachers must give students access to these documents for analysis, treatment and/or interpretation. The question is how the use of these sources is regulated. According to the Universal Declaration of Human Rights (UDHR, 1948) articles 26 §1 “Everyone has the right to education. (…)” and 27 §1 “Everyone has the right freely in the cultural life of the community, to enjoy the arts and share in scientific advancement and its benefits”. Also Protocol no 1 art. 2 of the European Convention on Human Rights (ECHR Prot. No 1, 1952) art. 2 §1 “No person shall be denied the right to education. (…)”. Education, however, needs to follow the development of science and technology. Content that is freely available online should be covered by lawful access (art. 3, Rosati 2021, pp. 27-28 and 45-46).

p-ISSN: 2184-044X e-ISSN: 2184-1489 ISBN: 978-989-35728-0-1 © 2024

3. Findings: Handling data – Analysis, interpretation, treatment

3.1. Transparency: Revealing errors

Considering word processors, we must admit that MS Word supports error-recognition the best with the Buttons Show/Hide and View Gridlines (Barnhill, 2017; Curts, 2017) (Figures 1–3). When these buttons are turned on, the non-printing characters, picture anchors, table borders, etc. become visible, they can be revealed in just one look. However, it is not widely accepted to share these findings. Unlike in industry, these errors remain hidden, even if they are recognized. In our time, it is a shame to reveal and share errors.

However, whether the non-printable characters are displayed on the screen or not, end-users' lack of knowledge, their ignorance and overconfidence (Kruger & Dunning, 1999; Staub & Kaynak, 2014; Gibbs et al., 2017) do not allow automated transparency tools to function properly. This entails that we need further tools to call attention to errors in digital texts and to their consequences, especially the soft-waste generated by negligent text management. Other tools of transparency, designed to reveal errors in digital texts, are ERM (Error Recognition Model) (Sebestyén et al., 2022) which includes the collection, presentation, analyses, correction, and proper formatting of word-processed documents (Csernoch, 2017), ANLITA (Nagy & Csernoch, 2023), and the entropy and sustainability rate of digital texts (Csernoch et al., 2022, 2023).

Figure 1. An erroneous Europass CV whose author claims that she has "excellent command of MS Office tools". The document itself proves that it is not so.

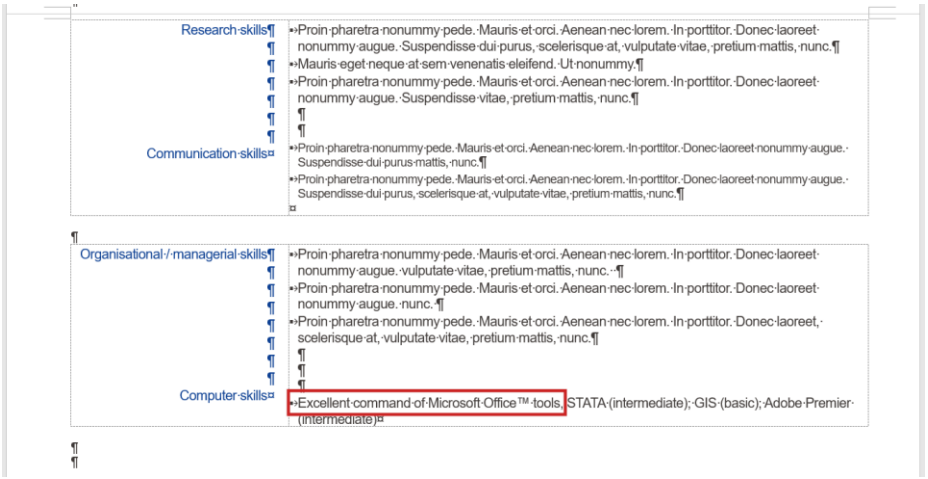
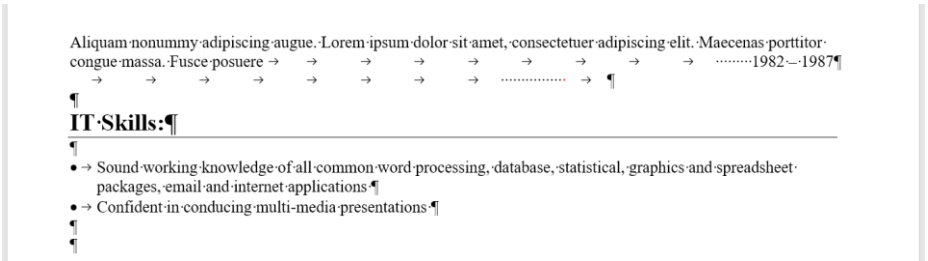


Figure 2. An erroneous individual CV whose author claims that he has "Sound working knowledge of all common word processing…". The document itself proves that it is not so.



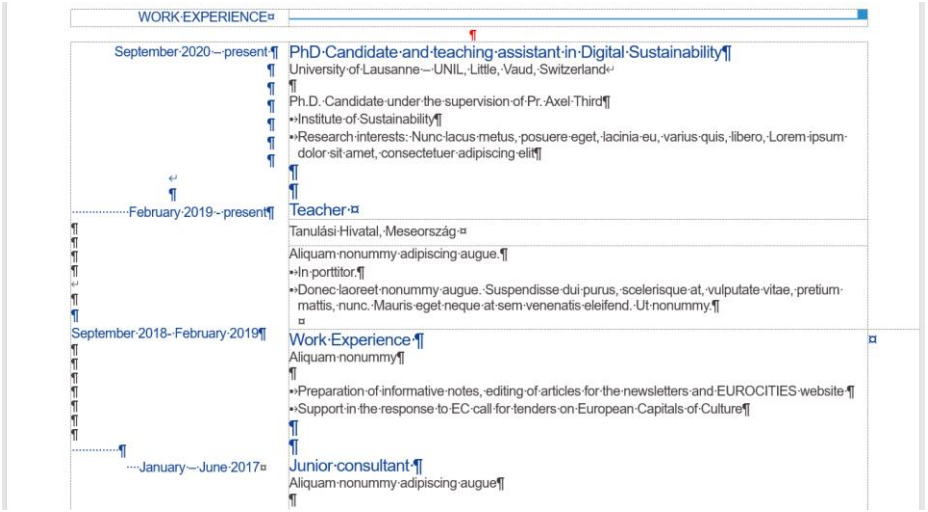
3.2. Digital resource vs. data – Europass CV

For the present paper to call attention to errors in word-processed document, one of the publicly available Europass CV templates are analysed (Bille & Kobosil, 2013) (Figure 3). The legal issues of using these documents are discussed in Section 4.

Templates, including CV templates by default can be considered digital resources, according to the DigCompEdu definition (Redecker, 2017), because officially they are immediately understandable to a human user. Considering the Europass CV Word templates (Bille & Kobosil, 2013) and the CVs created from these templates, it is obvious that the templates carry various errors and users make even more (e.g., Figure 1). Consequently, CV templates cannot be considered digital resources, since thorough analyses are required to reveal their true nature, which again proves the contradiction of the definition of the DigCompEdu (Redecker, 2017).

The Europass CV template carries various errors which should not reach the wide public in this form. One of the main errors is the structure of the template where the text is broken into several small tables open to change the sizes of the tables arbitrary. The original intention of breaking the text into small tables might be to indicate that complete tables should be copied to add entries. However, this message does not reach ignorant end-users.

Figure 3. The work experience section of a Europass CV where instead of copying small tables, the endless use of Enter (Paragraph Mark) and Space characters makes an erroneous and fragile document.



The negligent handling of pictures is another problem of the CV template. The formatting of the pictures is incorrect since their Wrap Text setting is Square instead of In line with text. This erroneous formatting leads to the arbitrary move of the pictures in the document. Several further errors are embedded in the CV template, including incorrect and contradictory font and paragraph formatting, styles, cell margins, etc., misplaced Paragraph Mark and Space characters, and attempts for their compensation.

4. Findings: Legal issues

4.1. Scientific research and freedom of expression

Art. 27 of the UDHR (1948) in §1-2 declares “Every-one 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.”

Scientific progress is important, and as shown above is a fundamental right. It is also necessary to note that scientific research itself is protected by the fundamental right of expression (ECourtHR, 2022b). 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. (…)” (art. 10 ECHR, 1950). Moreover, the International Covenant on Civil and Political Rights (ICCPR, 1976) recognize the “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 2 ICCPR, 1976). However, when exercising one’s freedom of expression potential violation of other fundamental rights need to be avoided.

4.2. Protection of property

The freedom of expression has been demonstrated in the previous section, however, §2 of article 27 of the UDHR declares that “Everyone has the right to the protection of the moral and material interests resulting from any scientific, literary, or artistic production of which he is the author” (UDHR, 1948).

Consequently, documents available online and so accessible to the public (Section 1.1) and protected by copyright (Bosher & Rosati, 2023) (Levola Hengelo BV v Smilde Foods BV (C-310/17, 13.11.2018)) are considered as (intellectual) property. Article 1 of Proto-col No 1 of the ECHR (ECHR Prot. No 1, 1952) provides a protection of property “Every natural or legal person is entitled to the peaceful enjoyment of his possessions. (…)”. (ECourtHR, 2022a) (Alleaume, 2009). Considering all these concerns, the question arises whether analysing documents in the classroom and publishing document sections in scientific papers create a violation of art. 1 of Protocol No 1 and other copyright rules (Bern Convention, 1979; Directive 2001/29/EC, 1997; Directive 2019/790, 2019) or not.

Eventual copyright violations must be clarified. The directive 2001/29/EC (1997) provides an exhaustive enumeration of exceptions and limitations to the reproduction right and the right of communication to the public in article 5 §3 letter A: “member States may provide for exceptions or limitations to the rights provided for in Articles 2 and 3 in the following cases: (a) use for the sole purpose of illustration for teaching or scientific research, as long as the source, including the author's name, is indicated, unless this turns out to be impossible and to the extent justified by the non-commercial purpose to be achieved;”.

Through the three-step test incorporated in article 9 §2 of the Bern Convention (1979), when in presence of limitation and exception it must be proven that (●) it is not overly broad, (●) does not rob right holders of a real or potential source of income that is substantive, and (●) does not do disproportional harm to the right holders (Sobrino-García, 2020).

In the present case, all the criteria are met listed above. Documents to demonstrate the numerous errors are presented objectively and accordingly to rules of computational sciences (Csernoch, 2017; Nagy & Csernoch, 2023), grammar, typology, etc. The content of the document needs to be part of the exercise so that students can better understand, relate to the topics presented, and carry out debugging. Important to ad, that the teachers should keep in mind for this reason the age and interests of the group they teach.

It is imperative that error-recognition in teaching text management, described in the present article, is integrated into education systems. Through this, a public interest arises which could potentially harm right holders. With this method, documents are analysed in classroom and results are published in scientific papers, the presentation of these texts must follow a proportional, “fair balanced” way (ECourtHR, 2022a). This requirement is fulfilled since the presentation limits to an objective demonstration of formatting and formatting related errors.

5. Conclusions

The present paper discusses how autonomation can be introduced in teaching text management and how digital resources can be used in this process.

A further concern of the paper is the legal issues of using digital sources available on the internet and in closed communities for classroom use and scientific purposes for the wider public. We found crucial that authentic sources must be presented in both cases to make content credible, interpretable, and correctable, and to avoid the errors of mass production systems that focus on both hardware and software tools. It is revealed that until the presented data pass the three-step test incorporated in article 9 §2 of the Bern Convention we can use them for educational and research purposes, without further restrictions on them. However, we must avoid the dissemination of personal data even if it is shared on the internet by the author of the original document. In this case, the data must be modified before handed over to students and/or colleagues.

References

Alleaume, C. (2009). À la croisée des sources de la Propriété Intellectuelle. L’article 27 de la Déclaration Universelle des Droits de l’homme. *Cahiers de La Recherche Sur Les Droits Fondamentaux*, *7*, 143-150. https://doi.org/10.4000/crdf.6707

Angeli, C., & Valanides, N. (Eds.) (2015). *Technological Pedagogical Content Knowledge: Exploring, developing, and assessing TPCK*. New York, NY: Springer.

Barnhill, S. (2017). *Word’s nonprinting formatting marks*. Nonprinting Formatting Marks. http://wordfaqs.ssbarnhill.com/NonprintChars.htm

Bille, H., & Kobosil, N. (2013, January 10). *Europass CV*. European Union. https://www.eea.europa.eu/about-us/jobs/application-documents/europass_cv_template.doc

Bosher, H., & Rosati, E. (Eds.) (2023). *Developments and directions in intellectual property law: 20 years of the IPKat*. Oxford University Press.

Csernoch, M. (2017). Thinking fast and slow in computer problem solving. *Journal of Software Engineering and Applications*, *10*(1), 11-40. https://doi.org/10.4236/jsea.2017.101002

Csernoch, M., Máth, J., & Nagy, T. (2022). The interpretation of graphical information in word processing. *Entropy*, *24*(10), 1492. https://doi.org/10.3390/e24101492

Csernoch, M., Nagy, K., & Nagy, T. (2023). The entropy of digital texts—the mathematical background of correctness. *Entropy*, *25*(2), 302. https://doi.org/10.3390/e25020302

Council of Europe. (1952). *Protocol 1 to the European Convention for the Protection of Human Rights and Fundamental Freedoms*. 20 March 1952, ETS 9 available at: https://www.refworld.org/docid/3ae6b38317.html, European Convention for the Protection of Human Rights and Fundamental Freedoms, as amended by Protocols Nos. 11 and 14

Curts, E. (2017, March 30). *How to show non-printing characters in google docs*. LaptrinhX. https://laptrinhx.com/how-to-show-non-printing-characters-in-google-docs-989631243/

European Court of Human Rights, Council of Europe. (1950). *European Convention on Human Rights and Fundamental Freedom*. Convention for the Protection of Human Rights and Fundamental Freedoms Rome, 4.11.1950, as amended by Protocols Nos. 11, 14 and 15, supplemented by Protocols Nos. 1, 4, 6, 7, 12, 13 and 16. https://www.echr.coe.int/documents/convention_eng.pdf

European Court of Human Rights. (2022a). *Guide on Article 1 of Protocol No. 1 to the European Convention on Human Rights. Protection of property*. Council of Europe. Updated on 31 August 2022, https://www.echr.coe.int/Documents/Guide_Art_1_Protocol_1_ENG.pdf

European Court of Human Rights. (2022b). *Guide on Article 10 of the European Convention on Human Rights*. Freedom of expression. Council of Europe. Updated: 31.08.2022. https://www.diva-portal.org/smash/get/diva2:1655124/FULLTEXT01.pdf

European Parliament. (1997). Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society. Procedure number: 1997/0359/COD. https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%3A32001L0029

European Parliament and of The Council. (2019). Directive (Eu) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC. https://eur-lex.europa.eu/eli/dir/2019/790/oj

General Assembly resolution 2200A (XXI) (1976). *International Covenant on Civil and Political Rights*. United Nations. Adopted and opened for signature, ratification and accession by General Assembly resolution 2200A (XXI) of 16 December 1966, entry into force 23 March 1976, in accordance with Article 49.

Gibbs, S., Moore, K., Steel, G., & McKinnon, A. (2017). The dunning-kruger effect in a workplace computing setting. *Computers in Human Behavior*, *72*, 589-595. https://doi.org/10.1016/j.chb.2016.12.084

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77*(6), 1121-1134. https://doi.org/10.1037/0022-3514.77.6.1121

McLaren, B. M., Adams, D., Durkin, K., Goguadze, G., Mayer, R. E., Rittle-Johnson, B., Sosnovsky, S., Isotani, S., & van Velsen, M. (2012). To err is human, to explain and correct is divine: A study of interactive erroneous examples with middle school math students. In A. Ravenscroft, S. Lindstedt, C. Delgado Kloos, & D. Hernández-Leo (Eds.), *Proceedings of the 7th European conference on Technology Enhanced Learning (ECTEL 2012)*. Lecture Notes in Computer Science 7563 (pp. 222-235). https://doi.org/10.1007/978-3-642-33263-0_18

Metcalfe, J. (2016). Learning from errors. *Annual Review of Psychology*, *68*(1), 465-489. https://doi.org/10.1146/annurev-psych-010416-044022

Nagy, K., & Csernoch, M. (2023). Pre-testing erroneous text-based documents: Logging end-user activities. *Frontiers in Education*, *7*. https://doi.org/10.3389/feduc.2022.958635

Redecker, C. (2017). *European Framework for the Digital Competence of Educators (DigCompEdu)*. (Y. Punie, Ed.). Publications Office of the European Union.

Rosati, E. (2021, August 26). *Copyright in the Digital Single Market: Article-by-article commentary to the provisions of directive 2019/790*. Oxford University Press.

Sebestyén, K., Csapó, G., Csernoch, M., & Aradi, B. (2022). Error recognition model: High-mathability end-user text management. *Acta Polytechnica Hungarica*, *19*(1), 151-170. https://doi.org/10.12700/aph.19.1.2022.1.10

Sobrino-García, I. (2020). Copyright in the scientific community. the limitations and exceptions in the European Union and spanish legal frameworks. *Publications*, *8*(2), 27. https://doi.org/10.3390/publications8020027

Staub, S., & Kaynak, R. (2014). Is an unskilled really unaware of it? *Procedia - Social and Behavioral Sciences*, *150*, 899-907. https://doi.org/10.1016/j.sbspro.2014.09.099

Tulis, M., Steuer, G., & Dresel, M. (2016). Learning from errors: A model of individual processes. *Frontline Learning Research*, *4*(2), 12-26. https://doi.org/10.14786/flr.v4i2.168

United Nations (1948). *The Universal declaration of human rights*. Paris, France.

World Intellectual Property Organization. (1979). *Berne Convention for the Protection of Literary and Artistic Works (as amended on September 28, 1979)*. https://www.wipo.int/wipolex/en/text/283698