

Why do we digitize books instead of knowledge?

Maja Turčić

University of Applied Sciences, Zagreb, Croatia

Abstract

Besides text and images, contemporary e-books can include multimedia, interactivity, pronunciation information and text-to-speech ready content. The current electronic publication standard (EPUB 3) provides all the technical solutions necessary for designing and displaying highly functional and content rich e-books, but the implementation is impaired by the e-reader industry, publishers and distributors.

Local and global book digitization has been an ongoing process for years, encompassing publications ranging from classical literature to current technical textbooks. Textual expression of knowledge is a remnant of the past but it is still dominant today, even though interactivity and video content are known to significantly enhance the learning process. However, multimedia and interactive content for e-books is not produced because it requires time, technical knowledge and specific tools.

Within the current publishing system, authors are expected to produce and design all the content for books on their own, but because of the extreme content diversity, it is unrealistic to expect that they really possess such knowledge and skills. Better collaboration among authors, publishers, distributors, the professional community and the related industry is necessary for knowledge digitization to function.

Key words: e-books, EPUB 3, digitization, knowledge

1 Introduction

The development of multimedia and interactive e-books has been an ongoing process for twenty years. The named media are struggling for dominance against mega-corporations, which are winning due to low prices and available e-readers, but also against the inheritance of the past embodied in the PDF format, whose text is unbreakable and universally available. Acceptance and practical use of the open EPUB 3 format is suspiciously slow, although technology is available, reading devices bought and users satisfied. Great efforts are invested into the digitization of already printed books. The information contained within them is then transferred to Google servers thus making millions of publications available to anyone with an Internet access. New pieces of work are also

created by corporations as well as by authors who opt for self-publishing. The education industry is also increasingly moving into digital spheres since users prefer interactive and animated materials. New media are more commonly present within *online* courses than in brick and mortar schools. If teachers create textbooks, it might be necessary to educate them on new formats which are available and appropriate for facilitating learning efficiency. The e-publishing industry does not prefer a particular format but trims its sails to the current disorder and short-term satisfaction of customers. EPUB 3 is based on web technologies and this should ensure simplicity of production and consumption, but current tools and the number of EPUB 3 e-books do not testify of it. Despite its advantages, the new e-book medium is not sufficiently present in e-bookshops or educational processes.

2 Redefining an e-book as a new medium

The way the definition of an e-book is perceived and applied impairs the development of information transition, primarily for educational purposes. For decades, an e-book has been perceived and used as a digital version of the “analogue” book, which is still current, although it was popularised by the invention of printing in the long past 15th century and it includes mostly textual and occasionally visual information [1].

Technological innovations which drastically change the means of use and creation are rare and do not threaten the existing industry but add value to the society and the economy (television has not destroyed the radio, nor has the Internet destroyed television). When a concept development relies on the previous version of the same idea, there is a possibility of the obstruction of the concept. In that case an innovation stops being innovative, and at best becomes a creative adaptation [2]. In this particular case, an e-book follows the development of devices used to consume it, the so-called e-readers. Amazon’s Kindle device (2007) presents a great progress with its *e-ink* technology enabling long-term reading owing to the low-power display and the comfort of reading on a screen with no backlight. The progress is best demonstrated by the fact that in 2011 Amazon had a greater sale of e-books than printed books [3] (although this economic leap could be interpreted by more favourable prices of e-books in comparison to prices of printed books). On the other hand, Kindle moves away from the universal idea of an e-book by presenting its own format *mobi*, as opposed to the first open standard format *Open eBook Publication Structure* (OEBPS).

Apple’s iPad resulted in the second big step towards accepting the new definition of an e-book. iPad is appropriate for reading because it is more practical than a laptop, and sufficiently larger than a mobile phone thus enabling a more pleasant reading experience. Besides supporting contents which had not been supported at any e-readers before,

Apple provided software for easy creation of multimedia e-books, but also in its own format *iBooks*. The situation changed in 2018 when it started giving its full support for the EPUB 3 format and terminated the development of its own proprietary format. Apple also offers high-quality software for reading the EPUB 3 format with the best in-built software for the accessibility for the blind and the visually impaired on all its gadgets.

The EPUB 3 format, as a huge development breakthrough after EPUB 2, conceptually redefines an e-book. Similarly to the radio, television and the Internet, an e-book is a new medium requiring a different approach to content creation. After five centuries of letters and images placed within a fixed frame of a page, the book loses its dimensions, and it offers the experience of listening, watching videos and animations, and enjoying completely accessible contents for all types of conditions and readers [4]. The content is adjusted to any screen size and operating system and removes the physical features of a book. The format supports the audio recording reproduction, but also the creation of an audio book with a complex system intended for the blind or the visually impaired readership, the value of which is increased by its application in educational literature [5]. Videos and animations are easily implemented owing to new Hypertext Markup Language (HTML5) elements which are integrated into the standard. Scripting support enables creation of interactive elements. An optimal e-book format implies an open standard format, searchability, adjustment to all screen sizes with an emphasis on readability, accessibility to persons with disabilities and support of all alphabetic characters [6].

The new specification 3.2 follows the latest web technologies and rejects its own ideas of support, and International Digital Publishing Forum (IDPF), the umbrella organisation of the EPUB format, with the DAISY specification for audio books, joins W3C and thus simplifies and facilitates the progress and support. The EDUPUB specification is also being prepared. It is intended for educational

purposes and emphasises the need to enable interactivity, accessibility, testing and social networking possibilities [7], but the success of the adaptation is difficult to predict at the moment.

However, there is a significant difference between the web and books, first of all in relation to ownership, rights, as well as durability. Distributing books, unlike web, requires distributors and libraries. Distributors decide on the features e-books must (are allowed to) have to be published. Limiting features are primarily the size (2GB Apple), but also the theft and the copy prevention system *Digital Rights Management* (DRM). The current Amazon policy stipulates that a user can read a book he/she has bought on six devices, and the user name cannot be inherited. In other words, e-books never become an actual property of a buyer. Publishers control the type of content to be published and the prices of books on the market. The most popular e-reader technologies are mainly owned by publishers (Apple, Amazon, Barnes&Noble) and thus they also control the support of “advanced” specification technologies. The interdependence of technology, publishers and distributors actively manipulates and obviously hampers the progress of e-books as a medium. A significant step forward and away from the existing book paradigm requires a breakthrough such as the one when Apple decided not to provide support to the Flash technology on its mobile technology operative system iOS, which paved the way to sudden progress and acceptance of open standards of HTML video and audio elements and the development of Cascading Style Sheets (CSS) animation [8].

3 Book digitization purpose and consequences

Books digitization began in 1971 with the well-known Project Gutenberg, which nowadays offers more than 18 000 publications in the electronic format. Project Gutenberg and similar projects aimed at preserving books

result in texts that can be searched, selected and copied. However, mass digitization of books that began with Google Books, and which many libraries have embraced, results in images of pages whose sole purpose is searching since they are illegible and cannot be downloaded locally. Google Books and Open Content Alliance aim to digitize all material printed so far but there are also more targeted projects of preserving rare books and making educational texts available [9]. In just a few steps, digitization has jeopardised safe earnings of big players of the industry, reduced the production price, simplified and facilitated the production and although it has decreased quality for the sake of quantity it has also created the golden age of digital products consumption. When it comes to books, digitization makes millions of copies available, and enables limitless, non-filtered self-publishing [10]. It seems there must be some quality in the mass production that is flooding the market. Amazon offers access to thousands of e-books and videos for only 12 dollars per month, Scribd boasts with millions of audio books, e-books and serial publications. Constant competition of big players results in a flood of cheap products making consumers confused.

The most beneficial effect of digitization is the availability of texts. Even in Croatia, which significantly lags behind America and Europe in implementation of e-books, in 2001 the first free electronic books platform BEK was launched. Only Croatian authors are available but although it offers a significantly larger number of PDF format publications, it also offers around 10 000 EPUB e-books and each year it records an increase in the number of users [11]. Since 2009, the web portal *eLektire.skole.hr* has been providing free access to required reading book titles for primary and high school in all formats. Assisted with omnipresent, inexpensive smart phones the named web portal makes a big library accessible to users even from the remotest villages. The new system involves entrepreneurs that create e-bookshops and e-libraries such as *digitalne-knjige.com*, *e-knjiga.hr* and *TookBook.com*, offering

e-books in the PDF and the EPUB format for a consideration. Only two e-library models are available in Croatia at present: the one in which borrowing an e-book is free of charge, implying that books are free, which makes these publications and authors worthless; the second model involves paying a fee, rendering books unavailable when the fee is no longer paid, which enables reimbursement to authors, publishers and bookshops, and although it does not precondition quality, it enables it. The unfavourable environment for e-books in Croatia arises primarily from an adverse policy and the 25% tax on e-books. This is reflected in statistics showing that in 2015 only 2% of Croats read e-books. The awareness and availability of e-books rose suddenly in 2016 when BEK recorded a 50% increase in the number of users [12]. Only in 2019 the tax on e-books was decreased to 5%, and now is the same to the tax on books.

4 Multimedia efficiency in e-textbooks

For centuries, texts have been the sole means of transferring knowledge. The process of forming thoughts through words is fluid and functional, which is substantiated by millions of books on various topics and with varied purposes. Works of art such as *Ana Karenina* shall always provide a more profound experience if read in its authentic, textual form.

Multimedia are, nowadays, specifically and abundantly used in creating e-learning material. However, the modes in which materials are presented, the way they are combined and ordered not to overload the cognitive processes is of key importance [13]. Also, the type of multimedia is an important factor in learning efficiency. Animation can either facilitate or hamper the learning process [14], but in general it can be claimed that animation, unlike static graphics, assists the learning process [15]. Depending on the topic, multimedia can significantly impact the learning efficiency and should be implemented in “textbooks” [16]. The number of *online* courses has an exponential growth and

students attending them actively show positive effects of the interactive learning [17]. Random incorporation of multimedia and interactive contents into e-textbooks shall not result in a higher motivation, acceleration of learning or better content comprehension. But, when describing processes that most people cannot visualise, it is useful to include an animation which will not impede the remaining textual content. For instance, learning anatomy using educational videos, mastering foreign languages by listening pronunciation, or accompanying physics content with interactive experiments certainly presents an enrichment and facilitation of the learning process and it could eventually oust *ex cathedra* lectures.

5 Problems of creating and distributing e-books

For those distributors who wish to please their customers, the process of making an e-book is unnecessarily complex and includes three formats. In addition to the traditionally produced PDF format which results from the prepress process and works well on all devices, the *Mobi* format is required due to the omnipresence of Kindle e-readers. Actually, persons who spend most time reading e-books use *e-ink* devices since their batteries are extremely durable and their screens without backlight do not tire eyes. However, the latest statistics show a decline in the sale of *e-ink* devices because of the lack of possibilities to display colour, animation, videos and other currently already standard digital media [18]. The third format available on the market is EPUB. Even seven years after the release of the EPUB 3 specification, the outdated EPUB 2 displaying only text and images is still prevalent. Most commonly EPUB 2 is sufficient for an e-book formatting, and publishers need to ensure accessibility on all devices, so the simpler and older version is safer. Videos are avoided in publishing due to the resulting size of the document. Optimization of EPUB e-books for all e-readers, either hardware or software, is impossible. There are more e-readers than web browsers and they are significantly different in their support of various formats. Tools used for formatting

e-books produce unpredictable versions which then have to be tested on a significant number of devices. The key tool, the EPUBCheck validator, which is the only means of checking the validity of EPUB e-books coding, is unreliable: various EPUB e-books occasionally pass the validation but do not display correctly while sometimes they do not pass validation but display everything. In such a complex and apparently non-functional ecosystem it is difficult to design, format and sell e-books. Who are those who create videos, animations, photographs and audio recordings for educational purposes and in which way can they capitalise their efforts? Contemporary digital knowledge is expected to be purchased at low or bargain prices while the production of multimedia e-books in 3 formats is in reality comprehensive and complex [19].

6 The environment required for EPUB 3 e-books acceptance and progress

Self-publishing is becoming a new standard in the publishing industry, which is testified by over a million of self-published books in 2017 [20]. Authors need to create interesting contents in order to please the audience faced with a wide range of offers. Most authors choose the Kindle platform for its user-friendliness and an easy access to customers via Amazon[21]. The EPUB format is an ecosystem designed to meet the needs of users. It is simple, flexible, based on web standards, ready for forming textbooks, fiction and serial publications. The publishing industry starts to accept this open format, but it is important to create a balance between richly equipped e-books and the readability on all devices and platforms. The EPUB standard, based on the web standards, implies constant evolution without which it would soon be outdated. Something similar happened to the Internet Explorer, which did not follow web trends and was consequently removed by force. [22] At the moment the created ecosystem does not meet the needs of authors. It implies knowledge of many different web standards that include HTML, XML and Scalable Vector Graphics (SVG) for contents, CSS for design,

JavaScript (JS) for interactivity, Pronunciation Lexicon Specification (PLS), Speech Synthesis Markup Language (SSML) and Synchronized Multimedia Integration Language (SMIL 3) for *text-to-speech*, and Accessible Rich Internet Applications (WAI ARIA) standards for accessibility. This makes the system incomprehensible to an average computer user, and even developers need to learn how standards function within the EPUB 3 format. Although layouting and/or e-books designing software, such as Adobe InDesign, iBooks Author and Sigil, provide for the creation of the EPUB 3 format, such e-books often do not get validated and are also not readily understood by an averagely IT literate person. Book authors first need to get to know potentials and limitations of a new medium in order to design the best version of an e-book for a specific situation and a specific audience. Including a video is important if it contributes to the content, but it is not necessary to include 50 video files lasting up to 50 hours. Space limitation, in addition to finding authentic and important information, is the biggest issue of the modern era, and it is true in relation to e-books as well. Developers need to learn about formats, but also e-readers which are different from web browsers and they need to test each specific possibility in order to make books available to most readers. Most readers have more than one device available, so it is not necessary to make contents available on each of them. It is to be expected that if a video cannot be reproduced at an *e-ink* reader the reader will access the same book or only the video content on the personal computer or a tablet, where it will certainly be reproduced. Publishers need to take into consideration the competition as well as the audience. It is not enough to offer the minimum, while at the same time the prices have to be reasonable. The education industry needs to have a privileged status and take the responsibility for creating the best possible materials under the auspice of the competent ministries. E-books are able to replace the traditional *ex cathedra* teaching and reach even the furthest regions of the Earth. Adequate materials available today can increase the effectiveness of learning and availability of knowledge.

7 Conclusion

An open, accessible and technologically adequate e-book, which conceptually differs from a printed book, is not the most popular format on the market due to the struggle of corporate policies as well as the chaotic industry of e-readers, which are not developed in line with the EPUB standard. When needed, books can be much more than texts and images and similarly to web pages need to accept videos, animations and interactivity as a common content.

A positive consequence of digitization is the distribution of knowledge enabled by an access to millions of publications. Via smart phones and the Internet, libraries are literally within arm's reach to almost everyone. When it comes to disadvantages, digitisation does not choose titles but floods the market with

old, new, good, bad and most commonly cheap texts and bewilders users with such a wide range of offer.

Adding animation and multimedia into educational contents needs to be well premeditated and done solely for the purpose of enriching the reading experience and facilitating the grasping of a specific topic. Such contents need to be created by teaching staff in cooperation with IT experts, and the industry surrounding the EPUB 3 format has to ensure adequate tools. The education industry needs to demand new and more accessible methods of contemporary transfer of knowledge. Readers, in an attempt to understand and access relevant information, seem to find their ways in the ecosystem of devices well and they readily forgive the industry its omission to provide support, which must not be an excuse not to create modern contents.

8 Bibliography

1. De Meester, B., De Nies, T., Ghaem Sigarchian, H., Vander Sande, M., Van Campen, J., Van Impe, B., ... & Van de Walle, R. (2014). A digital-first authoring environment for enriched e-books using epub 3. *Information Services & Use*, 34(3-4), p. 259-268.
2. Godin, B. (2014). Invention, diffusion and linear models of innovation: the contribution of anthropology to a conceptual framework. *Journal of Innovation Economics & Management*, 15(3), p. 11-37. DOI:10.3917/jie.015.0011
3. Chen, H., Hu, Y. J., & Smith, M. D. (2019). The impact of E-book distribution on print sales: analysis of a natural experiment. *Management Science*, 65(1), 19-31.
4. Garrish, M. (2011). What is EPUB 3?. "O'Reilly Media, Inc."
5. Ikuta, S., Nagano, S., Sato, E. T., Kasai, M., Ezoe, T., Mori, K., & Kaneko, C. (2019). Original Teaching Materials and School Activities with E-Books Containing Media Overlays. In *Handmade Teaching Materials for Students with Disabilities*. p. 76-110. IGI Global.
6. Turčić, M. & Pap, K. (2018). Dynamic mathematical layout in e-books, *Technical gazette*. 25(2).
7. International Digital Publishing Forum (2016). EPUB for Education, <http://www.idpf.org/epub/profiles/edu/spec/> [cited 2020 July 8]
8. Keizer, G., 2017. FAQ: How Apple, Google, Microsoft and Mozilla will eliminate Adobe Flash. *ComputerWorld*, July, 31.
9. Coyle, K., 2006. Mass digitization of books. *The Journal of Academic Librarianship*, 32(6), p.641-645.
10. Robinson, L. and Halle, D., 2002. Digitization, the Internet, and the Arts: eBay, Napster, SAG, and e-Books. *Qualitative sociology*, 25(3), p.359-383.
11. BEK - Besplatne elektroničke knjige, <https://elektronickeknjige.com/bek/> [cited 2020 July 8]
12. Zajović, Milena (June 1st 2015). U Hrvatskoj umire e-knjiga: Korisnici ih nisu navikli plaćati, a i ponuda je loša, *Večernji list.hr*, <https://www.vecernji.hr/kultura/u-hrvatskoj-umire-e-knjiga-korisnici-ih-nisu-navikli-placati-a-i-ponuda-je-losa-1008279> [cited: 2020 November 20]
13. Ying-Hua Guan. (2009). A Study on the Learning Efficiency of Multimedia-Presented, Computer-Based Science Information. *Journal of Educational Technology & Society*, 12(1), p. 62-72. Retrieved from <http://www.jstor.org/stable/jeductechsoci.12.1.62> [cited 2020 July 8]
14. Schnotz, W. and Rasch, T., 2005. Enabling, facilitating, and inhibiting effects of animations in multimedia learning: Why reduction of cognitive load can have negative results on learning. *Educational Technology Research and Development*, 53(3), p.47.
15. Berney, S. and Bétrancourt, M., 2016. Does animation enhance learning? A meta-analysis. *Computers & Education*, 101, p.150-167.
16. Iskandar, A., Rizal, M., Kurniasih, N., Sutiksno, D.U. and Purnomo, A., 2018, November. The Effects of Multimedia Learning on Students Achievement in Terms of Cognitive Test Results. In *Journal of Physics: Conference Series*. 1114(1), IOP Publishing. DOI: <https://doi.org/10.1088/1742-6596/1114/1/012019>
17. Durrington, V.A., Berryhill, A. and Swafford, J., 2006. Strategies for enhancing student interactivity in an online environment. *College teaching*, 54(1), p.190-193.
18. Derek Haines, Kindle Sales - The E-Reader Device is Dying A Rapid Death, justpublishingadvice.com, 28 September 2019. [cited 2020 July 8]
19. Laura Brady, Good Enough: A Meditation on the Past, Present and Future of EPUB, epubsecrets.com, January 18, 2018. [cited 2020 July 8]
20. Chi J, Lee J, Kim N, Choi J, Park S (2020) Secure and reliable blockchain-based eBook transaction system for self-published eBook trading. *PLOS ONE* 15(2): e0228418. <https://doi.org/10.1371/journal.pone.0228418>
21. Waldfogel, Joel. 2017. "How Digitization Has Created a Golden Age of Music, Movies, Books, and Television." *Journal of Economic Perspectives*, 31 (3): 195-214, DOI: 10.1257/jep.31.3.195
22. Bill Kasdorf, Key Issue - EPUB 3's Coming of Age, *Insight* 26 (2). p. 210-13, DOI: <http://doi.org/10.1629/2048-7754.06>