# EPUB, a standard,
# and its evolving adoption among e-book publishers

**Chloé Girard**,
*Paris-Nanterre University*

*Abstract*:
The first EPUB standard was set in 2007. Nowadays, although few studies have been made on this topic, EPUB has been adopted in quite different ways by French publishers, mainly depending on whether they belong to a group or are independent. By taking a quick look at this standard's history, its creation and evolution, we see the obstacles and eventual incentives that have shaped electronic publishing. EPUB is said to be an initially poor standard that is gradually drawing nearer to its target (the book). There is also talk abut the digital acculturation of professionals whom technology does not fascinate and about a still emerging market.

Since 2008, the National Book Center (Centre National du Livre, CNL) subsidizes publishers for the "*digitization and digital diffusion of already copyrighted documents*". In 2012, on the grounds that publishers were equipped with a production chain for making books in EPUB/XML, the CNL set 31 December 2011 as the deadline for the publication of works eligible for this subsidy. After this date, the subsidy would be lower or no longer available. In 2014, the production of "*EPUB according to the standards in effect*", till then optional, became a requirement. In April 2015, the deadline for the publication of the works eligible for a subsidy was pushed forward till 31 December 2014, "*a fixed, definitive date*". This deadline has not been reset since then. Can we draw the conclusion that, since 2014, French editors have widely adopted and use EPUB (which was standardized in 2007)?[1]

## Variable adoption rates depending on the category of publishers

According to a rapid calculation, less than 20% of new publications on paper in 2017 were produced in electronic format: the number of new books in electronic format was 22,720. From this figure, we have to subtract the 6624 books that were conversions of already copyrighted publications thanks to subsidies from the CNL in 2017.[2] This leaves us with 16,096 new electronic books in 2017 compared with 81,263 on paper, *i.e.*, 19.8%, But this statistic calls for refinement.

---

[1]This article, including any quotations from French sources, has been translated from French by Noal Mellott (Omaha Beach, France). The translation into English has, with the editor's approval, completed a few bibliographical references. All Web links have been consulted in April 2019

[2] FRENCH MINISTRY OF CULTURE (2018), "Économie du livre. Le secteur du livre: Chiffres-clés 2016-2017" (Paris: Direction Générale Des Médias et Des Industries Culturelles, Service du Livre et de la Lecture) March, Paris. Available at http://www.culture.gouv.fr/Thematiques/Livre-et-Lecture/Documentation/Publications/Chiffres-cles-du-secteur-du-livre/Chiffres-cles-du-secteur-du-livre-2016-2017.
CENTRE NATIONAL DU LIVRE (2018), *Bilan des aides 2017*, p. 208, available at http://www.centrenationaldulivre.fr/en/ressources/rapports_d_activites/.

The big French publishing houses have declared that they produce all new books in EPUB3 format. If we are cautious enough to take this declaration as referring only to "printed" books (thus excluding "picture" and "coffee table" books) and to subtract the number of new books in the latter category from new books in EPUB format, we obtain at most a rate of production of e-books for the remaining publishers of about 7% — regardless of the electronic formats used, since we do not know what is hiding behind the phrase "electronic book" in the studies conducted by the French Ministry of Culture and Communication.

We can thus draw the conclusion that publishers of a certain category have actually adopted the EPUB standard whereas the adoption rate is low or even zero among independent publishers. As we know, there is a multitude of small publishing houses in France. No study has yet been made of the actual adoption rate.


## A constantly evolving standard

EPUB was developed during the first decade of the new century to respond to the need to read documents on different sorts of screens. At the time, the PDF format produced a fixed layout poorly tailored to reading on small screens. In contrast, EPUB is a reflowable format. As happens with Web pages, the layout changes as a function of the reader device's parameters and of the reader's preferences, in particular as a function of the screen's dimensions and of the choice of a font and its size.

At the time, the first e-readers with e-ink (or e-paper) were coming out on the market.[3] These unconnected devices had low refresh rates; their technical strong point was the electronic paper screen. This coincidence of EPUB and e-readers was no accident. EPUB responded to the needs of e-readers. EPUB was a mini Internet site without any computer script. It accepted no actions such as drop-down menus or pop-ups, nor any elaborate formatting of the sort that constitutes the very art of typesetting and editorial work: hard spaces, the layout of elements in relation to each other, the separation of words at the end of line, widow and orphan lines, inserts or even full justification of the text. This format upset many a publisher and reader. The International Digital Publishing Forum (IDPF), a trade association which emerged in 2005 out of the Open EBook Forum, released it. Among its members, the IDPF counts publishing houses, Microsoft, Google, Apple, Adobe, and Barnes & Noble. It officially standardized the EPUB format in 2007.

Though of poor quality compared with HTML for the Web, the EPUB had the following advantages:
- It took account of the emerging e-reader industry's technical requirements.
- Since it is an open format, anyone can develop software for making or reading EPUB documents.
- By producing copies separate from the Web, it addresses the retail electronic book market, with its business model based on copies (like books on paper).
- It can be locked via digital rights management (DRM) so as to keep unauthorized copies from being made. This is also possible with PDF.
- Compared with websites, its offer corresponds to the "order" of actions related to books: reading, related browsing features, a full list of the resources for making a book.

The standing question was the variable (haphazard) page layout. During the 2010s, the first connected tablets, the equivalent of highly portable computers, came out. For magazines, picture or comic books, Apple specified for its iPAD the EPUB "fixed-layout". This set the height and width of an EPUB document down to the pixel for formatting the contents (text and images). Whereas EPUB, like HTML, had broken with the reference to printed matter, the EPUB fixed-layout came back to it. This format is often compared to PDF, with the disadvantages that arise on small screens: text cannot be

---

[3] Sony Librie in 2004, Bookeen's Cybook Gen3 in 2007.

adapted to the reader's preferences, since it sets the same display on all tablet screens. However it filled a need, and the IDPF standardized it in 2012 in EPUB3, which contains the specifications for Apple Fixed-Layout.[4]

The major players in the distribution of e-books, in particular Apple and Amazon, have strongly influenced EPUB and its market. Their integrated stores (Amazon/Kindle, Apple iBookstore/iPAD, FNAC/Kobo) accept no other format than EPUB — but an EPUB converted to their own closed formats (.ibooks, .mobi or .azw), which are encrypted and enhanced with proprietary data so that no other applications or readers can read their e-books. This has led to vertical device-content markets.[5] Lest they lose these big markets (29% of purchases in France), publishers have to produce an EPUB version that complies with both the IDPF's specifications and the required or recommended specifications of the giant firms in distribution — specifications that evolve constantly, changes about, for example, the size of images or the addition of *ad hoc* metadata.[6]

The EPUB norm has several versions, the most recent being EPUB 3.1, approved on 5 January 2017 by the EPUB Working Group Charter.[7] Reading devices have changed considerably: e-books are now mostly read on tablets, smartphones and laptop computers, the e-reader now ranking in fourth place.[8] The format now includes JavaScript, HTML5 and CSS3 (in other words, the latest Web standards). Furthermore, the on-screen display is ever more enhanced and graphic: text-to-speech, accessibility for persons with handicaps, etc. As the IDPF stated in an undated post, "*EPUB can, in principle, represent anything that you might deliver via an HTML5 website as long as the core content assets (HTML and CSS files, images and other media) can be determined in advance so that they can be stored in the .epub file*".[9]

EPUB and the Web are converging, as we see. Logically enough, the IDPF merged on 30 January 2017 with the W3C consortium (in charge of developing standards for the Web) and its Publishing Work Group.[10] The EPUB3 Community Group now handles questions related to the EPUB standard.[11] EPUB is destined to become a "*type of Web publication*",[12] a sort of HTML that accepts a technical terminology, an ordered semantics and a form of document type definition (DTD) suitable for "books". This convergence is to lead to a fourth version of EPUB, while maintaining EPUB3, since sales and production have been organized around it. At the time of writing, EPUB3 is still the current standard, but the specifications for EPUB3.2. are being finalized.[13]

[4] IDPF (2012), "EPUB 3 Fixed Layout Documents", 13 March, http://www.idpf.org/EPUB/fxl/.

[5] For two examples: AMAZON.FR AIDE "Service de documents personnels Kindle" at https://www.amazon.fr/gp/help/customer/display.html?nodeId=200767340; and "Apple Books Asset Guide 5.2.9" at https://help.apple.com/itc/booksassetguide/en.lproj/static.html.

[6] "Apple Books Asset Guide 5.2.9" available at https://help.apple.com/itc/booksassetguide/en.lpro.

[7] "EPUB 3.1. Recommended Specification 5 January 2017" available at http://www.idpf.org/epub/31/spec/epub-spec.html. EPUB Zone, "Understanding EPUB 3" at https://www.EPUBzone.org/EPUB-3-overview/understanding-EPUB-3/.

[8] SOCIÉTÉ FRANÇAISE DES INTÉRÊTS DES AUTEURS DE L'ÉCRIT (2018), "Baromètre sur les usages du livre numérique SOFIA/SNE/SGDL. La lecture de livres numériques: Des habitudes qui se confirment" available at https://www.sne.fr/actu/barometre-sur-les-usages-du-livre-numerique-sofiasnesgdl-2018/.

[9] EPUB ZONE, "What kind of content is EPUB useful for?" available at https://www.EPUBzone.org/EPUB-3-overview/what-kind-of-content-is-EPUB-useful-for/.

[10] Group created on 12 June 2017. *Cf.* W3C, "Publishing Working Group Charter" available at https://www.w3.org/2017/04/publ-wg-charter/.

[11] EPUB 3 Community Group: https://www.w3.org/publishing/groups/epub3-cg/; about the participants: https://www.w3.org/community/EPUB3/participants.

[12] KASDORF B. (2017) "The importance of EPUB and the need for EPUB 4" at https://w3c.github.io/publ-bg/docs/EPUB4_business_case.html. KASDORF B. (1998), "SGML and PDF: Why We Need Both", *Journal of Electronic Publishing*, *3*(4) available at https://doi.org/10.3998/3336451.0003.406.

[13] "EPUB Content Documents 3.2: Draft Community Group Report 27 March 2019" available at https://w3c.github.io/publ-EPUB-revision/EPUB32/spec/EPUB-contentdocs.html.

This overview of the EPUB standard's history suggests reasons why independent publishing houses have not adopted this format.

## Production, durability and maintenance

By 2010 and despite questions related to the layout (mostly settled by EPUB3), several publishers (by nature, not geeks) were hesitating about adopting the format since they did not know whether it would prove to be the "right" choice, *i.e.*, whether this format would last. Such a choice seemed adventuresome. Within a twelve-year period, EPUB has moved from versions 2 to 3 and soon 4. Upon release of EPUB3, the announcement was made that EPUB2 would be maintained; the two standards would evolve in parallel, respectively for "enhanced" and "homothetic" books. As we have seen, that did not happen. The current requirement is to produce EPUB3 format in all cases.

The catalog of EPUB2 books (now the biggest catalog) will, we imagine, soon become obsolescent. Some publishers are already investing in converting their catalogs to EPUB3. Likewise, we have doubts about whether how long EPUB3 will last. It will very probably be necessary to set aside funds for the conversion to EPUB4

The switch from printed to electronic books implies the passage from "storage" to "maintenance". As the experience acquired with XML/EPUB over the past twelve years has shown, maintenance is an imperative and, too, doable task. It entails adjusting typesetting and layout tools (whether externalized or not) and reproducing full catalogs by conversion to the latest standard. Publishers are concentrating on these technical points. Several compositors now produce the PDF, XML and EPUB formats for a price close to that of typesetting for PDF alone; and the code is of good quality (what did not use to be so). Nonetheless, the cost of maintenance has to be calculated. For the standard to be widely adopted, this cost might have to be covered by a subsidy; and the deadline for converting reluctant publishers' existing catalogs, extended.

The remaining question has to do with the switch in production processes. The drawbacks are not just economic: the time needed to find new partners, the eventual ejection of former partners, the adaptation of the publisher's tools for preparing data and metadata, a double verification of proofs (on paper and electronic), the move to electronic distribution, etc. Each of these task represents a substantial cost for small publishers, a cost that tends to be "human" (time and training). Meanwhile, the chances for a small company to be integrated in a big industrial group are declining. Even though the CNL has, since 2015, offered aid to "*independent publishers for the production of electronic books*", this offer has not met with success: only five publishers have benefitted from it.[14]

Does this mean that layout and typesetting have been (or tend to be) externalized? We observe different cases in the production of e-books. Some publishers have not made any changes in their production process apart from adding, downstream in the process, the recourse to specialized services for producing EPUB and XML out of PDF. Other publishers still deliver word-processing files (whether raw text or text enhanced with layout features) to multi-format compositors, most of whom seem to have found a way to cope with the switch from a graphic-centered approach toward an approach using tags to produce different on-screen results. Traditional compositors have had to adapt by becoming, in a way, webmasters; and they can now prove their mettle as trusted partners.

---

[14] On the CNL's website, *indépendants* figures in the menu but not in the description of this subsidy: On these subsidies, see https://www.centrenationaldulivre.fr/en/portail-de-demandes-d-aides/.

https://www.w3.org/publishing/groups/EPUB3-cg/

## A still uncertain market

In contrast with the technical obstacles to switching to EPUB, what factors favor doing so? We can still have doubts about the advantages of investing in the e-book market. In France, this market amounted to 0.5% at the start in 2009 and then rose to 5.7% in 2014 and 8.3% in 2016-2017.[15] These statistics were not broken down by category (EPUB, PDF, online texts and audiobooks).[16] In 2016, e-book readers overwhelmingly used the "*standard formats: PDF 43% and EPUB 28% (+10%)*".[17] Notice that PDF remained the favorite. Once again, the data were not broken down by format. In 2018, EPUB's share in the sales of French publishers was low, even very low.

## Conclusion

Caution is needed when discussing the adoption of the EPUB standard by French publishers. No detailed study exists for comparing specific formats, including other electronic formats. Such a study would have to take notice of the difference between types of publishers — and distinguish independents. The EPUB and XML formats entail switching technology and breaking with traditional work processes and with the formerly dominant "graphic" approach. The past twelve years have, we can conclude, lifted some doubts (about typesetting and layout) but reinforced others (about market trends, EPUB updates and maintenance needs). Meanwhile, the personnel in the book industry has learned about the electronic production of books. This experience of learning about code and computers is evidence that the full-fledged adoption of EPUB by independent publishers will take more time and require aid (new programs or the extension of existing programs) for preserving and updating existing e-book catalogs as the standard itself evolves.

---

[15] FRENCH MINISTRY OF CULTURE (2011), "Économie du livre. Le secteur du livre: Chiffres-clés 2009-2010" (Paris: Direction Générale Des Médias et Des Industries Culturelles, Service du Livre et de la Lecture) March, Paris. Available via http://www.culture.gouv.fr/content/download/19232/164647/version/1/file/Chiffres-cles_Livre_2009-2010.pdf.

[16] The market for audiobooks seems to be exploding: nearly two out of ten people in France have listened to an audiobook. SYNDICAT NATIONAL DE L'ÉDITION (2018) "Le livre audio, une nouvelle dimension au plaisir de lire" available at https://www.sne.fr/actu/le-livre-audio-une-nouvelle-dimension-au-plaisir-de-lire/. For a catalog of audio books, consult Lizzie: https://www.lizzie.audio/.

[17] SOCIÉTÉ FRANÇAISE DES INTÉRÊTS DES AUTEURS DE L'ÉCRIT (2017), "Baromètre sur les usages du livre numérique SOFIA/SNE/SGDL" available at https://www.sne.fr/actu/barometre-sur-les-usages-du-livre-numerique-sofiasnesgdl-2018/.