

From content edited for print page to reflowable EPUB 3 eBooks : building a methodology aiming at automating and sustaining the publishing of eBooks

Participant's interest

Publiwide has developed a suite of software to transform existing editorial content into reflowable eBooks. We are not interested in changing publishers and authors behaviors in regard to the authoring tool that they are currently using for creating print books. Based on that premise, the technology we have worked is syndicating the various source files published by content owners and provides toolset to streamline the production of EPUB 3 reflowable files.

With the aim of automating a large part of the process, we are seeking to learn more about any experience related to complex layout of eBooks with advanced work in CSS and HTML architecture. This concept is represented in our solutions by the "eBook template" which specifies the presentation of an eBook and the various possibilities of displaying the content according to the reading app capability (page swipe like in iBooks, content scroll like in Publiwide Reader and Vital Source). Considering the opportunity of dynamically formatting the content according to the reading app, we wish to offer to publishers scalable solutions to monetize their eBooks across the increasing number of EPUB 3 reading app.

Point of View

We believe that a vast majority of small to medium publishers expect to benefit from professional services and software tools that don't require any disruption of their current editorial workflow. In order for them to contribute to the eBook market, they need methodology to transform their content to whatever is required to comply with the Open Web Platform. Publishers are scared by the increasing number of market places, thus reading apps, supporting a slightly different interpretation of EPUB 3. We believe in increasing the automation of all the process liaised to the production but also the updating mechanism of eBook. This will be achieved by a much more comprehensive use of the open web platform components.

During two years of research and development, we had to elicit and tackle various requirements and problems:

- Enhanced user experience: the definition of techniques to transform static data into interactive and stimulating information.
- Digital publishing: the usage of open web technologies (for content and representation of content) to create assets that allows editor to publish attractive contents, respecting the requirements of their domain.
- Focus on the transformation process of preexisting data:
 - o Content: extraction and revamp of text, raster / vector images, formulas, footnotes tables, lists.
 - o Layout: transform to reflowable and fixed elements for hybrid layout definition.
 - o Styling: extraction of styling attributes as color, fonts (styles, size, ...), background patterns, etc.

- Automatic and supervised analysis: solutions for bridging the gap between printed books and information born for digital distribution.
- Data persistence, annotations and management: how to value the transformation of data, by abstracting its representation, from both content and form point of views.