W3C Workshop on Electronic Books and the Open Web Platform: Position Paper

Bill McCoy
International Digital Publishing Forum (IDPF)

Our Interest

The International Digital Publishing Forum is a non-profit trade and standards organization comprised of over 300 organization members from 38 countries. Our membership is drawn from the whole value chain of publishing including publishers, eBook retailers, software and device manufacturers, service providers, libraries, government agencies, educational institutions, and regional publishing associations. IDPF's mission is to foster global adoption of an open, accessible, interoperable digital publishing ecosystem

IDPF has since 2000 developed a standard file format (originally OEBPS, since 2007 called EPUB) for eBooks and other digital publications based on XHTML, CSS, and related Web Standards. Unlike pre-paginated formats such as PDF, EPUB is designed to adapt gracefully to different sized displays. With eBook adoption accelerating in recent years, EPUB has rapidly become the prevalent open standard format. For the last several years, all of the "Big Six" publishers in the United States have been delivering a single EPUB file per title to all of their distribution channels. Many eBook retailers and device/software manufacturers distribute EPUB files to consumers, including Apple iBooks, Barnes&Noble Nook, Kobo, Google Play Books, VitalSource, OverDrive, Sony Reader, and many others. EPUB is also used as an interchange format: publishers submit EPUB files to vendors such as Amazon, who convert EPUB into proprietary distribution formats, and web-based service providers like Safari Books Online, who ingest EPUB content into their cloud database and give consumers browser-based reading experiences.

The latest version of EPUB, EPUB 3.0 [1], was approved by the IDPF membership in October, 2011. EPUB 3 is based on HTML5, CSS3 modules and other elements of the modern Open Web Platform, enhancing layout options and accessibility and adding support for rich media, interactivity, and global languages.

IDPF has had an established liaison with W3C for several years and has collaborated particularly closely with the W3C CSS Working Group. IDPF is pleased to have contributed during the development of EPUB 3.0 to reinvigoration of work on global language support features in CSS such as vertical writing and ruby. IDPF looks forward to even closer future collaboration with W3C both in work towards future standardization specific to digital publications, and in helping to ensure that needs of our publishing industry focused membership are met in the evolution of the broader Open Web Platform for eBooks, websites, and native applications built with Web technologies.

Point of View

Publishers need to deliver digital book content as packaged eBooks, websites, and native applications built with Web technologies. It is not economically viable for publishers to create and manage many distinct versions of each title. Therefore it's critical for publishers to have a consistent platform that enables them to wherever possible deliver a single publication format that can be utilized everywhere and – even where specialization is called for - to reuse tools and components across these different modes of distribution.

IDPF made the decision to align EPUB 3 with HTML5 and related specifications - the Open Web Platform - based on the expectation that HTML5 and related specifications would become a universal runtime, beyond just for web page content rendered within a browser. With Flash and Silverlight fading as proprietary attempts at cross-platform runtimes, and Java increasingly focused on server solutions, and unusable by designers, the Open Web Platform is the only realistic option to achieve the long sought-after "write once, run anywhere" goal for rich interactive experiences.

Packaged eBooks (aka portable documents) represented via EPUB 3 and native applications built with web technologies are both typically supported via "headless" browser-engine-based runtimes, and both are generally self-contained and can be used both online and offline. However it is not clear whether the W3C as a whole has as yet fully embraced the notion that packaged apps and portable documents are truly peer-level use cases with browser-rendered web pages.

In particular, while it is encouraging that a System Applications Working Group has recently been formed [2], there is presently extreme fragmentation of different and non-interoperable approaches to native applications built with web technologies [3]. The Open Web Platform concept seems in need of additional specificity: it is said to include over 100 specifications [4] but these specifications do not always have a clear relation to each other and there is no definition from W3C of the overall architecture and what exactly comprises the ready-to-use Open Web Platform circa December, 2012. There have been efforts to more specifically profile collections of standards, most recently by the Core Mobile Web Platform Community Group [5], but there is no current activity on this that carries an official W3C imprimatur.

For eBooks the situation is arguably better as EPUB has been far more broadly adopted than any alternatives. However, there is still significant risk of fragmentation and divergence. In some respects the eBook ecosystem is today where the web was in the late nineties: we have vendors striving for complete lock-in, in some cases with an "embrace and extend" strategy based on EPUB, which leads to non-interoperable results like "works only in IE" websites back then. So we cannot get bogged down in overly elongated standards processes, or the open platform represented by EPUB could ultimately lose out to proprietary solutions.

Ensuring that packaged apps and documents are first-class in the Open Web Platform, along with websites, is not inconsistent with a preference for online solutions that connect users to the "corpus" of Internet content via its URL/REST-based architecture wherever possible. But not all experiences make sense to be wrapped in the "chrome" of a browser application, and offline and occasionally-connected usage is still important.

HTML5 came into being in large part to create an open alternative to proprietary solutions such as Adobe Flash and Microsoft Silverlight. Flash and Silverlight however were never just browser plug-ins, they were positioned by their respective vendors as universal runtimes, with desktop-class application instantiations (AIR and XAML) and document instantiations (FlashPaper and XPS). HTML5 has successfully "killed" Flash and Silverlight but has not yet fully matched these solutions in capabilities, particularly with respect to application and document use cases. And, in part due to single-source implementations, these platforms exhibited a significant degree of holistic architectural consistency, along with mature authoring tools. For the Open Web Platform to truly succeed as the universal open runtime these capability gaps must be expeditiously addressed while preserving interoperability and avoiding fragmentation of the platform.

At the same time we must keep in mind that, notwithstanding the growing importance of mobile apps and application runtimes, a core strength of the Web is declarative document representations. Fundamental content structure is an area of central interest to the publishing industry, an interest that transcends how content is packaged and delivered. And in helping educational publishing migrate to digital learning materials, handling content that is both rich and structured will be critical.

Suggestions

1. W3C should unequivocally declare that the Open Web Platform is a universal runtime: not just for websites running in browser applications but equally for packaged eBooks and native applications built with web technologies, for use online or offline.
2. The overall architecture of the Open Web Platform should be clearly communicated by W3C, including an over-arching framework for specific baseline profiles.
3. W3C should make it a highest priority to encourage interoperability of tools, libraries, and runtimes used to build native applications with web technologies. Continued fragmentation here risks disrupting the overall Open Web Platform vision.
4. IDPF and W3C should collaborate on a shared vision and roadmap, building on EPUB 3 as the standard packaged format, for the eBook (portable document) instantiation of the Open Web Platform. By the next major release of EPUB, maximum harmonization of EPUB with the overall platform should

be realized, with the expectation that the same web technology based tools, libraries and runtimes can be used across websites, eBooks, and native apps.

5. W3C should increase its focus on cross-cutting concerns that are applicable to multiple instantiations of the Open Web Platform. This includes security models, internationalization, device capabilities, accessibility, widget definitions, etc.. Wherever possible capabilities should be standardized for the whole platform not just for websites, apps, or documents. On the other hand, it should also be recognized that urgent priorities for one set of stakeholders may not always be considered mature enough for horizontal standardization and that to head off proprietary alternatives rapid progress in a particular vertical area is frequently needed. Therefore additional means to standardize rapidly at an interim level should be accommodated, that is more interoperable than e.g. vendor prefixes per specific implementation. The –epub- prefix approach utilized for certain CSS constructs in the EPUB 3.0 CSS Profile could potentially be generalized.

6. W3C and IDPF and other stakeholders should collaborate on declarative solutions for complex content such as infographics and assessments in education, including accessibility solutions.

References

[1] http://idpf.org/epub/30
[2] http://www.w3.org/2012/sysapps/
[3] http://www.i-programmer.info/professional-programmer/i-programmer/4982-the-disasterous-fragmentation-of-web-apps.html
[4] http://www.w3.org/QA/2011/01/100_specifications_for_the_ope.html
[5] http://coremob.github.com/coremob-2012/