Against the Grain

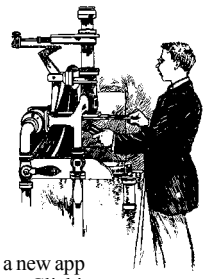
Volume 22 | Issue 5 Article 43

November 2010

Standards Column-Is the World Wide Web Dying? And Where Are the Standards for "Apps?"

Todd Carpenter
Niso, tcarpenter@niso.org

Is the World Wide Web Dying? And Where Are the Standards for "Apps?"



by **Todd Carpenter** (Managing Director, NISO, One North Charles Street, Suite 1905, Baltimore, MD 21201; Phone: 301-654-2512; Fax: 410-685-5278) <tcarpenter@niso.org> *www.niso.org*

The print copy of *Wired* magazine's September issue arrived in my mailbox with an eye-catching orange cover proclaiming the death of the Web. The feature article by **Chris Anderson** and **Michael Wolff** (*http://www.wired.com/magazine/2010/08/ff_webrip/*) points out with a colorful graphic that while we may be spending a great deal of time sharing information over the Internet, we are increasingly not using the World Wide Web as our primary interface. We are entering a world where devices, applications, and services are our entry point to content on the Internet.

I am probably a typical example of the behavior described by **Anderson**. Instead of reading the *New York Times* or *Wall Street Journal* in a browser, I have dedicated applications for those publications. I stream **Netflix** either through an application or via my **Wii**. **iTunes**, **LastFM**, and **Pandora** are my music portals, as well as where I stream many podcasts and radio shows. **Twitter**, **Facebook**, **LinkedIn** and **Skype**, where I carry on a fair amount of my communications, are all applications, not plain vanilla browser interfaces. Most, if not all of these, do have browser-based interfaces that I could use but they lack some of the functionality I have come to expect. Although, **Anderson's** article was pilloried in some tech circles for its misleading use of graphics (*http://www.boingboing.net/2010/08/17/is-the-web-really-de.html*), and overstating known trends (*http://techcrunch.com/2010/08/17/wired-web-dead/*), his article and post highlighted a growing problem with our interactions online, not just for users, but also for content creators, aggregators, and libraries.

Back in the mid to late1990s, development of online journal platforms was challenged by the need to test out the various browsers (*http://upload.wikimedia.org/wikipedia/commons/7/74/Timeline_of_web_browsers.svg*) to see how a site would be rendered and to ensure that the site functioned properly however users accessed it. In the early days of Web publishing, browser differences could make a site nearly unreadable on some of them. Testing on different versions of **Netscape**, **Internet Explorer**, **Mosaic**, or **Opera** was a critical component of pre-launch work to ensure that the coding was appropriate for the rendering. This is less the case now, although some variations remain.

Today, we're stepping back to those days of needing a proprietary software application and perhaps losing the interoperability we've come to take for granted with the Web. **Jonathan Zittrain** (*http://cyber.law.harvard.edu/people/jzittrain*) at the **Harvard Berkmen Center for Internet and Society** (*http://cyber.law.harvard.edu*) is one of those watching this trend and who decries the move away from open standards and integrated technology, which he argues drove the success of the Web.

If we are indeed moving to the "Age of the App" where Internet users have to interact with content via some interface that is not a browser, this will have significant implications for publishers. While I am a big fan of publication-specific apps, such as *Slate*, the *NY Times*, the *Wall St Journal*, *Wired* and others, not every publisher — indeed most publishers — are not in a position to create and maintain such an app. They'd also have to modify the app for the **iPad** platform, the **Android** platform, the **Blackberry** platform, various e-readers, etc. Plus there are all the devices that may develop next year or three years from now and all the different device's software upgrades that go on continuously. A figure quoted frequently earlier this year during the **American Association of Publishers/Professional Scholarly Publishing** meeting was that a good custom-built app could cost upwards of $50,000, not counting the cost of the post-release support and tweaking. A publisher's $50,000 development investment might have a shelf life of 12-18 months because of upgrades to the platform operating system that require an app upgrade or complete redesign. If building one $50,000 application is on the verge of being too expensive for your organization, building three or four is simply not an option.

The cost alone would be a big impediment for many smaller publishers. An even more critical problem is that the publisher now has an application that works on selected devices but not on others, resulting in only partial penetration within the community for the publisher. The user is also affected by having to install (and possibly purchase) a different app for every publication and launch a new app when switching publications. Clicking on links within the publication can launch yet another app (or ironically, a Web browser window). The library community is further challenged by serving diverse communities only some of whom may access a portion of the licensed content.

Operating system changes, platform dependencies, and user demands for increased functionality have been problems since the advent of electronic publishing. But the World Wide Web's success, especially as an information distribution platform, was due to its ability to circumvent most of these issues and that ability was due to the underlying standards infrastructure. The era we seem to be entering is taking us back to those earlier problems, multiplied by a much larger variety of devices to support. In an App world, the only standards are the de facto proprietary platform standards used by each device. Although there is some advocacy for standards, such as EPUB for eBooks, most eBooks are still issued in the proprietary format of each e-reader usually wrapped by some form of DRM, or the EPUB formatted publication is overlaid with the publisher's navigation app. From a user perspective, interoperability is even more critical than ever, because few people have only one device and they need to be able to move their content between their smartphone and their laptop, or their PDA and their organization's file server. This is exactly the kind of interoperability that requires the use of common standards, not proprietary applications.

Smaller publishers will likely have to partner with aggregators to deliver their content, much as they did with pooling resources for Web-based distribution platforms like **HighWire**, **Project Muse**, or **BioOne**. As yet such aggregators have not launched device specific applications. For the moment only larger publishers are venturing into the app space, such as the **American Institute of Physics** with their **iResearch iPhone App** (*http://scitation.aip.org/labs/10_15_09_iresearch_iphone_app*) released last year or the **Nature Publishing Group** (*http://itunes.apple.com/us/app/nature-com/id349659422?mt=8*) and **Public Library of Science (PLOS)** (*http://itunes.apple.com/us/app/plos-medicine/id362137769?mt=8*), each with multiple apps distributed through the **iTunes** store. Highlighting the underlying problem, though, is the fact that all of these applications are for the **Apple iPhone** or **iPad**, not for other platforms. Although **OCLC** has allowed its **WorldCat** data to be served up via third-party applications on a range of platforms, **OCLC** itself has also only developed for the **Apple** suite of products.

And where are libraries in this new app world? With ever-shrinking budgets, libraries can't afford to manage a digital collection with multiple proprietary versions of each content item and all the apps required to run them. If a library chooses (or is forced through budget constraints) to "standardize" on one or a few devices and platforms, they are then limiting the availability of content to what has been developed for those platforms. Just like smaller publishers, libraries will likely need to work with one or more aggregators to ensure access to all the desired content — when or if such aggregators are available at an affordable price. The preservation issues will also become even more complicated than they currently are in the browser-based environment, where libraries are still struggling with how to ensure preservation of content. As if preservation of digital content alone were not difficult enough, there is ample proof of how difficult

continued on page 18

Managing Our Collections ...

from page 1

collections given the complexities of campus stakeholder preferences. Sensitivity to user needs and ability to deploy strong rationales for decision-making can help leaders navigate difficult choices.

This ***Against the Grain*** issue focuses on managing print collections, but the truth is that each of the profiled initiatives is fundamentally about library strategy and services. In an environment of constrained resources, libraries strive to serve user needs with new formats and innovative support roles, find mission alignment with their parent organizations / funding bodies, and avoid deviating from the vital shared value of preservation. Finding the right balance for print collections is imperative to planning a strategy for the library to meet user needs in a changing environment.

Pelikan's Antidisambiguation

from page 14

have at hand, and if our "desktop" and everything we've left there, comes with it — well, that'll be a bunch of steps further toward the kind of environment many have been envisioning for a long time.

So let's all take a look at **Blio**. Regard it not as an app that runs on a **Windows** machine, but recognize it as the next step toward a uniform, multi-platform environment that goes where you go — and that isn't necessarily or automatically run by either of those twin gorillas, **Apple** or **Amazon**.

Google's a pretty big gorilla too. And **Microsoft** — a fair-sized gorilla itself — hasn't died off — not by half.

So I guess we're in for quite a show here. For myself, I'm going to grab some popcorn, a root beer, and enjoy all that emerges…

Heading West: Circling the Wagons to Ensure Preservation and Access

by **Emily Stambaugh** (Shared Print Manager, California Digital Library, WEST Assistant Project Manager) <emily.stambaugh@ucop.edu>

Research libraries have inherited a legacy of print duplication; duplication that made sense in its time to ensure institutional competitiveness. But a network-wide shortage of storage space requires us to reduce the physical footprint of retrospective collections. Research libraries seek ways to make informed decisions about what to preserve and what to withdraw. The recent growth in last copy agreements suggest there is real momentum in the community to find collaborative solutions.[1] But taken together, these efforts do not reach the scale that is needed to address the systemic and long-term shortage of space to house physical collections. Among the factors that have hampered such efforts, are: the absence of business models, organizational structures, collection decision-making models, disclosure systems, and incentives to create and sustain trusted archives. Large-scale collection consolidation has real operational costs that surpass existing consortial capabilities. A network level (regional, national, international) solution is required. Research libraries and consortia in the western United States have prepared a business model and operational structure for a **Western Regional Storage Trust (WEST)** which is designed to support network level archive creation services to preserve the scholarly record, provide access, when needed, and manage reallocation of space.

About Aggregate Print Journal Collections

Print journal archives are ideal candidates for space reclamation for reasons that are well-known; large amounts of shelf space can be reclaimed with a relatively small number of titles (and decisions about those titles). To put the size of the aggregate print journal collections in perspective, there are about 4.18 million print serials in **WorldCat** and the average number of libraries that hold a title is about nine. At the high end of the duplication spectrum are roughly 10,000 titles in **Portico** and **JSTOR** with average holdings of 250 and 600 libraries, respectively.[2] While titles in **Portico** and **JSTOR** are the usual suspects for collaboration, there is clearly a need for collaboration on other electronically held titles and on titles published only in print. As much as 40% of the refereed scholarly journal literature is not available in electronic format. Some 56% of peer-reviewed history journals are published in print-only format. By contrast, almost 80% of the refereed medical journal literature is available online.[3] There is an economic sweet spot for consolidating print collections, and it can be found where duplication is highest and where holdings can be compared in semi-automated ways for ready decision-making. The extent of possible candidates may be great enough to remedy library and storage facility space problems without dipping into more costly monograph deselection projects or more risky restrictions on collection growth.

In the western region of the United States, an initial analysis of print journals held by thirteen research libraries and their storage facilities revealed at least 60,000 commonly held journal "families" (current + previous titles of a journal). About 30,000 are held by 3 or more institutions in the region and about 17,000 by 5 or more (up to 21 copies). These duplication rates are probably understated at the title level, as a significant number of records supplied for analysis could not be meaningfully compared due to lack of match points (ISSNs). Further analysis is underway to compare regional rates of overlap network (national, international) level overlap.

Duplication Level	# Copies	# Journal Families (current + previous titles)	%
High	5 or more	17,233	28%
Moderate	3 or 4	13,381	22%
Low	1 or 2	29,966	49%
Total		**60,580**	

Table 1: Levels of Print Duplication within WEST Planning Libraries

On Collaboration Scale

The scale of collaboration requires careful consideration: state, regional, national? Creating archives at a certain pace has real operational costs and requires dedicated staff trained in project management and validation. In 2009, the **University of California** Libraries considered going it alone with a consortial archiving service that would serve the ten **UC** campuses. Experiments were conducted

continued on page 20

Is the World Wide Web Dying?

from page 16

preservation of content is in a world dependent on hardware and software integration.

Without open standards and open platforms, building applications to reach the end user will be ever more expensive, the availability of content could become limited by the choice of a device, and the cross-publisher and cross-platform linking we have come to depend on could break down. If we are indeed moving away from the era of the Web and toward one where the application is king, we need to start thinking about and advocating for the standards that will make the new world as accessible and interoperable as the one we're leaving.