September 2006

If It's Not Digital It Does Not Exist: The Future of Science and Technology Collection Development

Johann van Reenen

University of New Mexico, jreenen@unm.edu

Follow this and additional works at: <https://docs.lib.purdue.edu/atg>

Part of the [Library and Information Science Commons](#)

Recommended Citation

Reenen, Johann van (2006) "If It's Not Digital It Does Not Exist: The Future of Science and Technology Collection Development," *Against the Grain*: Vol. 18: Iss. 4, Article 7.

DOI: <https://doi.org/10.7771/2380-176X.4914>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

If It's Not Digital It Does Not Exist: The Future of Science and Technology Collection Development

by **Johann van Reenen** (Associate Professor, Assistant Dean for Research & Instruction Services, University Libraries, University of New Mexico, Albuquerque, NM; Phone: 505-277-8310) <jreenen@unm.edu>

Emergent behaviors and information systems will change collection development irrevocably. We can attempt to halt them (but they are unstoppable). We can attempt to change their results (although disruptive agents are unpredictable) we can capitalize (our only real option) on these emergent systems. These should be considered seriously — we live in times where what seems undoable, political suicide, or too expensive one day may be our saving grace the next.

As Wilson and Hayward (1999, not verbatim) says:

"The increase in the degree of [electronic] connectedness constitutes the greatest change in life today. For it is not just that there is more change than ever before, but the inter-connectedness has changed the nature of change itself, making it more unplannable and unpredictable, more abrupt and dynamic than

Pictured above is Johann van Reenen.

Convenient, Communitarian ... from page 18

dia help explain the text, make the argument, or verify the claims an author is making in her book, including it will become commonplace.

It will be interesting to see what sort of new creations authors develop with **Sophie** once it

it has ever been or than our traditional organizational systems can handle."

From the publishers view. Many of the author's ideas coincide with the Wittenberg (2006) article:

"While we have been busy attending conferences, workshops, and seminars on every possible aspect of scholarly communication, information technology, digital libraries, and e-publishing, students have been quietly revolutionizing the discovery and use of information. Their behavior, undertaken without consultation or attendance at formal academic events, urgently forces those of us in scholarly publishing to confront some fundamental questions about our organizations, jobs, and assumptions about our work." (Wittenberg 2006)

I have also been reminded (Thank you to **Linda Lewis at UNM**) of the **Janus Conference** of 2005 and the Six Key Challenges for Collection Development in Research Libraries they arrived at (<http://www.library.cornell.edu/janusconference/>).

Being in the Customer's Space

Our ability to serve information to a variety of devices is crucial. Our customer's space is more and more electronic and found in the handheld device of a Net-Generation student. "Few of our OPACs can handle searching from say a Blackberry!" Whose job is it to arrange technology development that will allow this? For services such as **Google** and **Yahoo!** the question is moot. They are in the customer's space and have beaten us to it!

is available. Needless to say, it has the potential to change our notion of the book. If it lives up to the publicity it will certainly be an important tool that will lead to our being able to provide collections that will meet our new undergraduates' expectations. Perhaps these will be in time to meet the **Rookwood** children's expectations by the time they are undergrads. 🐼

Endnotes

1. **Schiesel, Seth.** "Welcome to the Dollhouse: Barbie? Who's Barbie? Young girls are playing electronic house and redefining the audience for video games." *New York Times* 7 May, 2006: Section 2, page 10.
2. **Ugoretz, Joseph.** "Three Stars and a Chili Pepper: Social Software, Folksonomy, and User reviews in a College Context." In *Academic Commons*, posted June 9, 2006.
3. **Wittenberg, Kate.** "Beyond Google: What's Next for Publishers?" *The Chronicle of Higher Education* 16 June, 2006: B20.
4. <http://network.nature.com/boston>.
5. **Vershow, Ben.** "The Book Is Reading You." *Publishers Weekly*. 253.23 (2006) p. 70.
6. **Visel, Dan.** "The Institute for the Future of the Book presents Sophie: A quick introduction to the future of reading & writing." <http://www.futureofthebook.org/sophie/SophieIntro.pdf>. Viewed July 3, 2006.

Libraries need to:

- Set up electronic document delivery, e.g., 80 % same day email delivery. Services such as **RAPID** come close.
- Extend the pool of available monographs through a delivery system such as **Prospecter**. The **Colorado Library Alliance** has over seven million items in the combined catalogues that can reach a customer's desk in about the same time it takes her/him to go to the library and get it themselves
- Develop **Amazon** look-alike catalogs that present search results for owned and non-owned items with rich annotations and pointers. Provide options for delivery to home or office.
- Default ILL requests for items two years or newer to an on-the-spot (library credit card?) order as discussed in previous issues of **ATG**. This monograph will be used at least once, which is more than we can say about a large percentage of the books we purchase. Alternatively, partner with say, **Amazon** and **Powells**.
- Purchase electronic book services rather than eBooks themselves, especially in areas that use monographs in "bits and pieces" to check facts or refresh knowledge, e.g., computer and engineering sciences.
- Purchase more materials that engage users such as educational gaming tools and games. Then build information content for these.
- Provide alternative ways to deliver content, e.g., pod-casts to deliver instruction and training.

From Berlin to DC — Open Access

The global O.A movement will influence CD decisions as more content becomes available.

I believe we are nearing a tipping point when academic (including published peer reviewed) content will explode on the Web.

How credible is the OA movement? There is a growing record of accomplishments, culminating, for US libraries, in the **FRPAA** bill before **Congress**. The list includes the **Budapest Open Access Initiative** (December 2001), the **Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities** (October 2003 see: <http://www.zim.mpg.de/openaccess-berlin/>), the **Bethesda Declaration on Open Access Publishing** (2003), and the **National Institutes of Health Rule** (2005) that "requested" grantees put copies of their published articles in the agency's own online repository, <http://www.pubmedcentral.nih.gov/>, **PubMed Central**. This should

continued on page 24

If It's Not Digital from page 22

be strengthened by the US Federal Research Public Access Act of 2006 (FRPAA) (http://cornyn.senate.gov/doc_archive/05-02-2006_COE06461_xml.pdf).

A recent **European Commission Report** (April 2006) recommends open access to all publicly financed research. (See: http://ec.europa.eu/research/science-society/pdf/scientific-publication-study_en.pdf). The US is a late-comer, however, a May 2006 **Harris Interactive Poll** showed strong support (82%) from the general public, i.e., the taxpayers! (82%). (See: "Americans support free access to research" at: http://online.wsj.com/article_email/SB114893698047965609-MyQjAxMDE2NDMAMTkzMzE2Wj.html).

Six countries are considering national OA policies.

Is There a Viable Financial Model for OA Publishing?

The best analysis is in a report from JISC, the UK's **Joint Information Systems Committee**, (www.jisc.ac.uk) describing models that could underpin user services developed over repositories in the UK. The report is at: http://www.jisc.ac.uk/index.cfm?name=programme_digital_repositories.

Financial models are of concern as some calculations indicate that high-output research universities might pay more in author fees for OA than they do for journal subscriptions. The June **SPARC Open Access Newsletter** (<http://www.earlham.edu/~peters/fos/newsletter/06-02-06.htm>) provides more insight.

What is the Alternative?

Bergstrom and McAfee compiled data on **Journal Cost-Effectiveness Website** (www.journalprices.com). The average 2004 price of 1,756 journals from not-for-profit publishers was \$442.46, while the average price of 3,138 journals from for-profit publishers was \$1,418.61. The average for-profit journal is over three times as expensive as the not-for-profit. Their data (see: <http://www.hss.caltech.edu/~mcafee/Journal/Summary.pdf>) on publishers with a sampled output of 50 titles or more shows that **John Wiley and Sons** has the highest average cost of \$2,472.07 across 188 titles; **Elsevier** followed at \$1,926.76 across 1007 titles, and **Academic Press**, until purchased by **Elsevier** a fairly reasonable commercial publisher, is now third at \$1,478.75 over 158 titles. **Springer** rounded out the top four with 441 titles averaging \$1,449.69. Their new purchase, **Kluwer**, had 553 titles that averaged \$1,318.46.

Journals (including packages) are overpriced. There really is no choice for researchers and librarians but to collaborate on alternative publishing systems based on emerging ideas including open access, altruistic col-

lections based on social software, and other yet unknown ways to communicate discovery.

From Local to International Collections Development

More and more libraries and related organizations are banding together to create subject, geographic, and community-of-interest portals with digital collections that "belong" to the group.

Just a few examples from my own organization:

- **Searchable Ornithological Research Archive or SORA** at <http://elibrary.unm.edu/sora/>. An OA archive of the world's most important Ornithology journals. The Biology faculty negotiated agreements with the publishing societies. Imagine each academic library taking on one subject area to archive association-published journals.
- **Online Archive of New Mexico** at <http://elibrary.unm.edu/oanm/>. This is a single, integrated source for finding aids to archival collections in New Mexico.
- **Western Waters Digital Library** at <http://westernwaters.org>. This is an information resource for the Western United States' most precious resource; a multi-State aggregated collection using **CONTENTdm** and headquartered at the **University of Utah**.

UNM's Institutional Repository holds, as examples:

- **Meetings and proceeding.** For instance we host the minutes of the **University's Board of Regents** which they are required to make public
- **Whitepapers, opinion pieces, grant preparation work.** These are important as these ideas can stimulate other work even if a grant is not funded. It may also lead to new partners for resubmission of such a grant
- **Local journal or other publication series.** For instance, a professor in the Department of Economics at UNM, with the help of his graduate students, started a portal for the **Nepal Studies Center** (see: <http://nepalstudycenter.unm.edu/index.htm>) with a very easy and affordable journal publishing system based on our **Dspace** implementation. They currently have two start-up journals:

— **Himalayan Journal of Development and Democracy** (see: <http://nepalstudycenter.unm.edu/journals/hjdd.htm>)

— **Liberal Democracy Nepal Bulletin** (see: <http://nepalstudycenter.unm.edu/journals/LDNB/ldnb.htm>)

- **Association documents.** We are in negotiations to become the official repository for the **American Indian Planning Association** which has unique planning documents relating to developments in Tribal areas.

- **Research data.** For instance we have a large collection of **Scanning Electron Microscope (SEM)** images of cave microbiology (more about this below) collected from **Karst** systems (<http://en.wikipedia.org/wiki/Karst>)

Listen to what science and technology librarians hear from their customers — and **MAKE IT SO!**

- **If it's not digital it does not exist for me** — Librarians hear this from, especially, physical and engineering researchers and students. Natural science researchers are more conservative. Rather than arguing this point, strive to make it so with behind-the-scenes CD and digitization.
- **Everything I need is free on the Internet** — Rather than arguing this point, strive to make it so. Help customers be better users of **Google Scholar** and look for joint-venture opportunities with commercial players. Consortia could begin to work less as buyers clubs and more as innovation clubs. One example could be negotiating with **Google, Amazon, Yahoo!**, and other information competitors the development of services that integrate searching for information with directions to local free catalogues (public and university libraries), local book stores, (including directions through **Google Maps**), commercial bibliographic databases, and online purchasing options. The metadata to do this already exists for the most part.

The "big deal" came in like a lion but may leave with a whimper. It has not affected the market place enough to contain costs. We have seen lower annual increase rates and better consortial arrangements that acknowledge cheaper rates for each member as membership in a particular agreement grows. In any other market this would have been inevitable — in our business it took concerted efforts on all fronts to get where we are.

- **I can get everything I need from my colleagues** — Again, help make this long standing tradition more efficient by introducing customers to collaborative social software systems such as **Connotea** (<http://www.connotea.org/>). At my library we are working on two grants that include the use of social software to create collaboratories based on subject archives, wikis, social tagging, et cetera.
- **If it's electronic do we really own it or can it go away?** — This comes from unusually perceptive faculty and from the HSS areas in particular. Fortunately there are successful efforts underway (**Portico**) and functioning (**LOCKKS** and **CLOCKKS**) to begin to address perpetual ownership and archiving. CIDs need to ensure that what they buy is compliant with such services and that their libraries are participating in electronic preservation initiatives. Libraries already do a good job of only signing licenses that have perpetual access clauses.

continued on page 26

- **The customer as selector and cataloguer** — The “Living Web” (also called **Web 2.0**), describes the explosion of Web services based on collective intelligence (Levy and Stone 2006) and opens new ways of engaging customers and “entering” into their electronic workspaces. Community-built reference sources such as **Wikipedia** and **Flickr** are replacing reference tools not only for Net generation users but for librarians themselves.

What does this city look like, how do I get there, and what can I do there, what will it cost? Such questions can be answered by such services. The information is current and based on real experiences of the social network members constructing them. Libraries point out the weaknesses, but users want the “live” aspects and the community evaluations — maybe this is the new credibility? For instance, **Steward Butterfield** co-founder of **Flickr**, a photo-sharing and community-tagging site, says it’s the “eyes of the world” as quoted in **Levy and Stone**:

“When news happens anywhere in the world, it’s common for the first photographs not to be sent via news wire, but posted to a **Flickr** site.” For example, three people posted photos of the September 2004 bombing of the Australian Embassy in Jakarta before news persons could get there.

Should selectors suggest free sources for quick reference and as content for mash-ups to push to the customer’s space? Should they then stop buying the *Encyclopedia Britannica*, et cetera? I believe we are getting there fast. I myself have rarely opened an encyclopedia in over two years but use, e.g., **Wikipedia**, compulsively to check facts and learn about terminology and ideas I encounter for the first time. I have yet to be disappointed with incorrect or overly biased information, especially when a particular entry provides a good reference list. It is an excellent starting point; even for students. The informal “peer review” between “expert” contributors are fierce and based on not losing Web-face and their reputation as an expert. It presents academic problems for students if they quote **Wikipedia** and do no further in-depth research. This is a challenge for faculty and librarians *Chronicle of Higher Education’s* “Wired Campus” blog (<http://chronicle.com/wiredcampus/article/1328/wikipedia-founder-discourages-academic-use-of-his-creation>). (Permanent link to this article “**Wikipedia** Founder Discourages Academic Use of His Creation” at <http://chronicle.com/wiredcampus/index.php?id=1328>).

This is reality in a world of user choices and it behooves us to make these tools useful in an academic context. Market forces strongly support such services rather than libraries as information sources! **Bradley Horowitz** says of his recent purchase of **Flickr** for **Yahoo!**:

“With less than ten people [**Flickr**] had millions of users organizing that content for them [using social tagging], tens of

thousands of users distributing that across the Internet, and thousands of people not on the payroll actually building the thing” (Levy and Stone 2006).

Big Science and the Library

One of the most daunting challenges is maximizing the use of large data sets sitting around campuses long after research projects are completed. This data could be the source of future discovery if warehoused and curated correctly. The politics of who owns it and who/how/when it can be re-used is daunting but the library is probably the most likely organization to do this well. At my university the Library, the VP for Research and Economic Development, and the High Performance Computing Center are working together on this issue. The University Libraries are determined to be part of the solution, and preferably the lead organization.

Interdisciplinarity is having a significant impact. Like many academic libraries, we were expected to provide a collection to support a new inter/multi-disciplinary Nanosciences program just when we were planning a large journal cancellation. It would be another nail in the library’s coffin had we said no — it is after all a competitive strategy for the university whose mission we support. We used the discussions as a teachable moment to create opportunities for the group to investigate questions raised in “What do science and technology librarians hear from their customers?” above. We found that, in fact, some things were “free on the Internet” and could be integrated (selected, if you will) into our locally developed “Journal One-Search” service (<http://elibrary.unm.edu/jlsearch.php>). That does not negate due diligence to select what we could afford, such as described by **LaBonte** (2005):

“Citation analysis was used to determine if the **Sciences-Engineering Library** at the **University of California at Santa Barbara** is meeting the needs of an interdisciplinary group of 60 faculty members at the new **California NanoSystems Institute**. The latest three publications of each faculty member (published within the last two years) were analyzed in two ways using the Science Citation Index: 1) the journals they were published in, and 2) the journals where cited articles were published. The results indicate that the library subscribes to 98% of these journals. This information is useful to map the citation patterns of a new interdisciplinary field and can be used for future collection management decisions.”

Trends and Data

Data and trend analysis is critical. The usefulness of approval plans in science is doubtful. It is more efficient to let the customer tell us through focus groups, behind-the-scenes electronic data captures, direct requests. Approval plans were declining in science and technology collections by the mid-1980s. A study presented at the 1988 Annual Conference of the **American Society for Engineering Education** shows that 60% of STM books were selected directly and 29% by approval plans (Osorio 1988).

More recently, academic library Websites shows low use of these plans, for instance at the **University of Illinois U/C** (see: <http://www.library.uiuc.edu/acq/FY04apstats.htm>) approval books were:

- Under 4 % of purchases in Biology, Chemistry, Physics, and Statistics
- 1 % or less, in biotechnology, mathematics, and natural history
- 10-15 %, in all of Engineering (most likely standing orders and proceedings).

It is much lower in my organization, where most monographs are purchased based on a user-input system.

Statistical data is essential. Using **COUNTER** compliant vendors helps, as do services such as **ScholarlyStats** that track usage of large bundled services/e-collections.

However, the complete picture only becomes apparent when data from all sources are integrated (**LibQual+**, Customer surveys, focus groups, Web logs, circulation statistics, journal cost-per-use data). At **UNM** we created a position for a Data and Trends Analyst. Unfortunately, it was hard to fill and soon vacated. We will try again but in the meantime are investing in an in-house relational database to log instruction, reference, and other service and collection statistics.

The Human Side of Collection Development

Librarians need to understand and accept the importance of speed, ambiguity, risk, and complexity in molding decisions. Exposure to national trends, best practices at other libraries (e.g., **ACRL’s** online seminar, “Electronic Collection Development for the Academic e-Library” http://www.ala.org/ala/acrl/acrlprof/tools/ecollection_development, June 29-July 22, 2006), workshops that challenge the status quo and deals with strategies five years and beyond (see for instance <http://www.aisti.org/events/index.php>) and internal reorganization ensure that they keep up. It is important not to underestimate the value of the good librarian-detective and to deploy their skills to really make information “free” and findable. We should also capitalize on the credibility librarians have as certifiers of good information.

Conclusion

It is critical that we reorganize our libraries and set new expectations that include higher level skills in information technology. HR systems that prevents “group think” are also critical. At **UNM**, during a major reorganization in October 2005, we realized that selectors need to report unambiguously in the subject area where they collect, i.e., Science and Engineering selectors report to the Director of the Science and Engineering branch. It is becoming impossible to disentangle reference, instruction, library electronic service provision, outreach, and building local digital content from collection development.

Customers do not care how you obtain needed information (paid or free) or how hard you have to work behind the scenes to make it

continued on page 28

Collection Development & SURVIVAL in the "Mostly Digital" Library

by **Dr. Michael Stoller** (Director, Collections & Research Services, New York University Libraries, 70 Washington Square South, New York NY 10012; Phone: 212-998-2566)
<michael.stoller@nyu.edu>

As researchers have seen their access to remote materials increase, local collections have lost much of their significance. We have a unique opportunity to share, an opportunity that we have, on the whole, lamentably ignored. The moment has not entirely passed, and we should still take advantage of this unique chance. But, as paper collections decline in importance and digital resources move to the center, we'll face a new challenge that will dwarf the old resistance to collaboration that came from institutional pride.¹

The "Mostly Digital" Library

We are facing the dawn of the "mostly digital" library, a world in which the vast majority of primary and secondary materials on which scholarship rests are in digital form. It's "mostly digital," because some components seem unlikely candidates for digitization any time soon. The vast archival holdings of our special collections, millions of linear feet of correspondence, manuscripts and realia, while they will certainly be the objects of selective conversion and the source of wonderful digital exhibits, don't at present suggest a viable business model for mass digitization. And area studies may lag behind, though Latin America, the Middle East and even Africa may surprise us and turn to digital publishing more quickly than we imagine. But the paper vestiges of the journal world will surely become digital before the end of the decade, and, at some time very soon, that last great bastion of paper — monographic publishing —

will find viable business plan(s) for digital delivery or face extinction in a world of 500-copy print runs and remaindered back lists. Even textbooks and mass market publishing may soon convert, as students rebel against the outlandish cost of the former and a viable reading platform emerges for the latter. **Google, Yahoo!, Microsoft** and a host of entrepreneurs we've yet to imagine will digitize our legacy collections. The large, paper-based collections that have defined research libraries for centuries may not vanish, but they'll be little used, and the scholar of the next generation, like many scholars already, will spend most of her time at a computer.

Does this mean that the library, if not already "overwhelmingly a virtual destination," as **Jerry Campbell** has said, soon will be?² Absolutely. Does this mean, as many writers seem to think, that collection development is fading away like **Marx's** state, to survive only as license management. Absolutely not. The latter does not follow naturally from the former for a simple but often curiously ignored reason: these digital collections will not be free. Libraries will still need to decide how and where to spend their money, and, if we're not all to bankrupt the institutions we serve, we'll need to be able to shape those collections carefully, as we always have, and to exercise far more control over what we buy than the world of digital publishing has offered us thus far.

One thing we can say with some certainty:

in the mostly digital world, the local collection will not have the meaning that it had in the paper age. It is a privilege to browse the shelves of a great research library. I think it always will be. The scholar exiled at a small college with a tiny library lived a harsher research life than colleagues at a great research university with a rich library. But today a large part of that great research collection isn't on the browsable shelves. It's miles away in a remote storage facility, browsable only in the catalog. Just as accessible, however, are the collections of dozens of other research libraries in a world where interlibrary loan is fast and getting faster and remote collections are nearly as accessible to the small college professor as to his colleagues at the research university.

But there is a problem with this democratic vision of our new library world: it applies only to the paper collections, and, in the mostly digital library, those paper collections won't matter very much. What will matter is a scholar's ability to search the literary holdings of *Early English Books Online*, the journals of *Science Direct* or the government documents of the digital *U.S. Serial Set*. Scholars will want to search the full text libraries being built by **Google, Microsoft** and others, libraries that will surely not come free. But, unlike paper books, these digital resources, as the library world is structured right now, are not an interlibrary loan away. They're licensed to the institutions that buy them, and their use is almost entirely limited to affiliates of those institutions. In the "mostly digital" library, as it's evolving right now, that scholar at a small, rural college is back in the same researcher's exile that typified such places in the 1970s.

Reinventing Collection Development

This is exactly why collection development needs to be at the center of our thoughts, not a forgotten remnant of the days when we bought our collections one book at a time. As we wrestle with how to build this mostly digital future, we can't abandon collection development. Instead, we must reinvent it. It's not a local enterprise anymore. It's not about matching the book to the faculty member. It's about making sure that we have a truly national collection, that scholarship and the raw materials that are its sources — materials that will be mostly digital — are accessible to researchers everywhere.

So, let's talk a little about the reality of the mostly digital library. First, as anyone who's watched it grow in the last decade knows, it won't be free. Open access will be a part of the picture, at least in the sciences. But even that won't be free. If true open access models emerge, and scholars pay publication fees or

continued on page 30

If It's Not Digital from page 26

useful and to integrate it with their way-of-working. Do not waste time explaining. We "just" need to supply what they need quickly, profes-

sionally, and on their desktop in friendly and, if possible, interesting ways. This will require resources (obtained from stopping some of the things we do now that do not support the above), investments in flexible but standards-based and Web-based IT systems, and a re-aligned workforce.

References

1. **Bergstrom, Ted** and **Preston McAfee**. Journal Cost-Effectiveness [Website] "search engine to find internationally-published journals and rank them by price per article or citation." See: www.journalprices.com.
2. **Osorio, Nestor L.** (1988) Approval Plan Materials and Direct Orders: Differences and Similarities. In **Grayson, Lawrence P.** and **Biedenbach, Joseph M.**, Eds. Proceedings American Society for Engineering Education, 1988 Annual Conference: Beginning a New Era in Engineering Education 4, pp. 1729-1735, Portland, Oregon (USA).
3. **LaBonte, Kristen B.** (2005) Citation Analysis: A Method for Collection Development for a Rapidly Developing Field. *Issues in Science & Technology Librarianship* no. 43 (Summer 2005) p. 1 at: <http://www.isl.org/05-summer/refereed.html>.
4. **Levy, Steven** and **Brad Stone**. (2006) The new wisdom of the Web. *Newsweek*, April 3, 2006, p. 47 - 50, 52 - 53.
5. **Wilson and Hayward**. (1999) Winning through chaos—part 2. *Credit Control*, 20 (9/10): 35-39).
6. **Wittenberg, Kate**. (2006) "Beyond Google: What next for publishing." *Chronicle of Higher Education*; Section: *The Chronicle Review*. Volume 52, Issue 41, Page B20.