Chain reactions:
Archives and the value chain in chemistry's scientific communications

NRC Chemical Sciences Roundtable Workshop, October 26, 2004
Anna Gold, MIT Libraries

> Numerous scientists have pointed out the tragic irony that, right at the historical moment when we have the technologies to permit worldwide availability and distributed processing of scientific data and their concomitant promise for broadening collaboration and accelerating the pace and depth of discovery, we are busy locking up that data and slapping legal restrictions on transfer.[1]

Outline:
Introduction
1 – The problem(s) of archiving
2 - What the present enables, what the future could hold
3- Opportunities to lose: your stake as chemical scientists
4 - A concluding vision: A Greater Chain of Being

Introduction:

We have heard a great deal about problems of expensive scientific journals; in this last session we are focusing on some solutions to those problems. In order to see our way out of the present cost crisis, we need to broaden our view beyond the value chain associated with scientific *journals*: *journals* are part of a much larger value chain of scientific *communication*. This communication occurs before, during, inside, outside, and after the journal article is published. This suggests that we ought to reframe the question posed in this workshop, and ask ourselves not about the "needs" of chemical and chemical engineering journals, but rather examine the communication needs of chemists and chemical engineers. Is the publishing segment of the cycle serving the ends of the cycle as a whole? Could that segment serve the whole better than it does today? This analysis suggests that there are three critical solutions to pursue: creative commons; open access; and library-based archiving. Of these three solutions, open access to the literature of chemistry is a key to your ability to redistribute, reuse, add value to, mine, explore, and archive the record of your system of science, without the constraints of paper-based media, commercial ownership, institutional wealth, or fragmentation by publisher.

> "(A)ll scholarship is implicitly a negotiation with, an interpretation of, and a contribution to the archive."[2]

1. The problem(s) of archiving:

In a time of expanding vertical integration by scientific publishers, the problem of archiving has become a growing concern. The problem has multiple facets.

The first is *reliability* over time: if MIT cancels print receipts of *Tetrahedron Letters* in 2005, how can we ensure that the MIT scholar in 2064 will still be able to access issues from 2010? Can we rely on publishers to be responsible for archiving the print record of scientific communication and discovery? There are major publishers – societies as well as commercial

[1] creativecommons.org/projects/science/proposal

[2] Marlene Manoff, "Theories of the archive from across the disciplines," portal: Libraries and the Academy, v. 4 n. 1 (2004), p. 13.

entities that have had to turn to libraries and archives to recover missing content needed in the course of digitizing the very content they have published. That should tell us something.

Solutions: Cooperation among societies and libraries
LOCKSS is an approach to electronic journal archiving in which libraries and publishers cooperate to ensure minimal distributed access to redundant copies of electronic files. This model suggests there is potential for societies and libraries in particular to collaborate in assuring a long-term archive.[3]

The second is *context*: the scholarly record is an interlinked system of communications. Libraries and archives have a primary mission to preserve the scholarly record. Their interest is not limited to preserving the product line of a single publisher, but with integrated access across many publishers.

Solutions: Interoperability and federation
Context is critical to a meaningful archive. Many publishers are providing richer context by opening their content for federated searching via web search engines, although full text access to this content is still unavailable except via subscription. Linking services, enabled by standards such as DOI and CrossRef, also provide context and are highly valued by researchers. Interoperable and federated access is critical to providing full and accurate context; for this reason, cross-journal search capabilities such as Elsevier's ScienceDirect, or their forthcoming SCOPUS service, encompassing the output of only single publisher (plus open access journals), provide no substitute.

A third is *affordability*: Many libraries even at the wealthiest research universities are unable to sustain local collections of research journals during bad budget years. The last few years have seen waves of print cancellations at these libraries, and also waves of electronic cancellations.[4] For journals that are not reasonably priced, libraries will be unable to sustain long-term archival collections.[5]

Solutions: Open access and cost sharing
Open access is a promising solution to ensure access to primary research content, in the absence of publisher-added value. Cost sharing through interlibrary lending, consortial purchases, and cooperative archiving, may provide some respite as well.

A fourth is *intellectual property*. The problems of affordability are also problems of intellectual property. Concerning the new digital publisher archives, one author writes,

> "Journals and newspapers are beginning to realise that the Internet gives them the ability to sell their product many times over. Yesterday's newspaper is no longer today's fish-and-chip wrapping,

[3] For example, all BioMed Central Open Access journals address the LOCKSS (Lots of Copies Keep Stuff Safe) <lockss.stanford.edu> through permanent archiving arrangements with PubMed Central <pubmedcentral.gov>, at the University of Potsdam <www.uni-potsdam.de/over/homegd.htm> in Germany, at INIST <www.inist.fr/index_en.php> in France and in e-Depot <www.kb.nl/>, the National Library of the Netherlands' digital archive.

[4] Among many others, ACS President Charles P. Casey acknowledges these costs, proposing that "scientists…exert pressure on commercial publishers" (*Chemical & Engineering News*, v. 82 n. 4, January 5, 2004, p. 2-5).

[5] "Certain publishers exert pressure on libraries to abandon the archiving of print journals. By making it prohibitively expensive for libraries to maintain paper copies, commercial publishers are altering the shape of the historical record. In the new digital marketplace, publishers like Elsevier are having as great an influence on the scholarly archive as the academic community that produces it." Manoff, op cit., pp. 12-13.

> it is an archive. And, as time passes, access to those archives will become more necessary and higher charges will be exacted."[6]

The present model of permissions and costs threatens the loss of archival capacity and redundancy throughout the scientific world[7]. When costs of electronic plus print soar beyond the ability of libraries to sustain both, they will choose electronic access (because it is what you scientists prefer). They will no longer own, but rather only license, access to journals. They will lose the ability, because publishers prohibit this, to archive the digital content that they license. They will become completely reliant on publishers' ability to maintain these digital archives themselves. They will be vulnerable to publishers' changing business models for providing long-term access to digital archives.

Solutions: Creative Commons and open access

Lawrence Lessig[8] describes how "commercial culture," defined as "that part of our culture that is produced and sold or produced to be sold,"[9] has engulfed noncommercial culture. This has happened not only in artistic, but also scientific communication, which is fundamentally "noncommercial." It's not that scientific *discoveries* don't have commercial potential, but that is different. But with the Internet, scientific communication, like "the ordinary ways in which ordinary individuals shared and transformed their culture"[10], has fallen into the "reach of the regulation of the law," with the consequence that "we are less and less a free culture, more and more a permission culture."[11] By asserting the right to manage post-publication access, publishers threaten the ability of scientists to add to their own culture by accessing it, both now and over time. Some publishers argue that holding copyright prevents abuse of author copyright, and supports reuse, by providing a single point of contact for those who wish to reuse any content they publish. Enter Creative Commons,[12] a package of well-crafted legal agreements that allow authors to both *control* and *enable* reuse of their work.[13] The important point is that open access does not mean, as Rudy Baum of the ACS has recently asserted, "socialized" science (or the loss of copyright protections for the work).[14]

Fifth, there are *technical problems*. Digital archiving is still in its infancy. Unlike paper archives, digital archives ask us to rely on formats that are not, yet, "preservation" formats. Because of the technical unknowns, the economic costs of digital archiving are also unknown. The recurring costs of digital archives, including media migration, metadata, and storage, or ones for which we have little data or experience.

---

[6] (See "Who Owns John Sutherland?" *London Review of Books*, v. 21 n. 1 January 7, 1999, www.lrb.co.uk/v21/n01/suth01_.html

[7] Interview with Derk Haank in *Information Today*: "Copyright transfer is a very efficient way of making sure that we have the commercial rights to exploit the articles."

[8] Lessig, *Free Culture* (2004).

[9] Lessig, Ibid., p. 7.

[10] Lessig, Ibid., p. 8.

[11] Lessig, Ibid., p. 8.

[12] Creative Commons web site: creativecommons.org/

[13] See Creative Commons web site for a Science Commons, at creativecommons.org/projects/science/proposal

[14] Rudy Baum, "Socialized Science," *Chemical & Engineering News*, v. 82 n. 38, p. 7, September 20, 2004.

Solutions: Collaboration and Research
There are several initiatives by organizations with archival missions that together are addressing all of these problems. One is collaborative ejournal archiving of core publications by libraries and library organizations. JSTOR is an example of that approach, but it is unclear whether its cost-sharing model is sustainable. Another is the development of federated approaches to digital archives, exemplified by DSpace; DSpace is premised on open access, and has yet to resolve the long-term technical costs of archiving. A third is robust funding for research on digital archiving: federal agencies like the National Archives, the Library of Congress, and the National Science Foundation are soliciting proposals for research, in the expectation that inroads can be made on the social as well as technical challenges of digital archiving.[15] There are urgent needs for research to address what to archive (selection), how to manage rights over the long term (ownership), how to ensure the sustained funding needed to maintain the archive over time, how to provide appropriate access, and how to ensure interoperability with other archives.

Finally, there are problems of *longevity*. If properly governed and managed, archives help insure that both journals and associated data will exist in the future, whereas more informal and ad hoc modes of data management do not.

Solutions: partnerships among libraries, technologists, scientists. The survival of archives over long periods of time may require multilateral commitments from strong parties with independent governance and economic foundations.

> "The challenge of preserving digital information and guaranteeing permanent access to it can only be addressed successfully by realizing a long-standing and close co-operation of three key-players: leading memory institutions (national libraries and archives), main producers of information (publishers and public agencies), and, last but not least, leading IT companies. The development of the e-Depot at the KB together with the science publisher Elsevier and IT-company IBM is a good example of such a co-operation."[16]

2. What the present enables, what our future could hold

These problems suggest that it would be a mistake to rely on publishers to provide our future archives, and argue further for open access to the archive. Publishers don't need the archive to sustain their operations. Indeed, in internet-era businesses, content tends to zero price, with direct links to consumers, while a new source of value is "service-enhanced customization." In the publishing world, where content itself is not scarce, there is greater value in services such as indexing, filtering, personalizing, and recommender systems. There is profit to be had by publishers who add such value. A number of publishers have already taken up this challenge,

[15] See the DIGARCH call for proposals by NSF, 2004, www.nsf.gov/pubs/2004/nsf04592/nsf04592.htm. DIGARCH is a collaboration between NSF and LC's NDIPP program (National Digital Information Infrastructure and Preservation Program), whose vision "is to ensure access over time to a rich body of digital content through the establishment of a national network of committed partners, collaborating in a digital preservation architecture with defined roles and responsibilities." National Digital Information Infrastructure and Preservation Program, *Preserving Our Digital Heritage: Plan for the National Digital Information Infrastructure and Preservation Program, A Collaborative Initiative of the Library of Congress* (Washington, D.C.: Library of Congress, 2002), pg. 65; available in print and at www.digitalpreservation.gov.

[16] www.rlg.org/en/page.php?Page_ID=17068 about e-Depot's work with Oxford University Press. See *RLG DigiNews*, Volume 8, Number 2, "Treasuring the Digital Records of Science: Archiving E-Journals at the Koninklijke Bibliotheek:" A related article: www.researchinformation.info/risepoct04archiving.html

calling it a "hybrid" model.[17] And there is plenty of evidence that even the most elite publishers could survive very well on payments for value-added content:

> A significant fraction of *Nature*'s £30m revenue is spent to commission and produce the non-research-article content of the journal (e.g. News & Views articles, book reviews, commentaries, editorials etc.) This non-research content would continue to drive healthy print and online subscription revenue, even if the research articles were made freely accessible online. Since the non-research content (the front-matter) is far more widely read than the research articles themselves, it is far from clear whether making the research articles Open Access would have any negative impact on subscription revenue. In fact, the opposite can be argued.[18]

We are a long way from answering the question of who pays how much for open access, but we have a lot of choices.[19] The point remains that the core value of the content comes not from the publisher but from the scientific community and from those in the community who read, use, and cite it:

> Open-access publishers say the important thing is not who nominally pays the costs - in the end it all comes from the same, usually publicly-funded, pot - but how the research is used. 'Everyone should have access to publicly-funded science,' PLoS's Harold Varmus told the Select Committee. The impact of open access comes from the combination of instant access and searchability. 'It's about a fundamental change in the way scientific findings are recorded and used.[20]

Future productivity and the commons: The truly exciting thing is that this debate over the excessive cost of scientific communication is already shifting toward the related issues of control (copyright, licensing) and use. In this context, a new agenda to enable and shape the future of scientific communication is now being set. The impact of Open Access and related initiatives (e.g., institutional and national repositories[21], Sherpa's Romeo ratings[22] of publisher copyright policies, the Directory of Open Access Journals[23]) is already considerable. The agenda is now shifting to an even more exciting opportunity, to enhance the future productivity of science. Creative Commons is an emerging leader in this, through their Science Commons proposal to create licenses that can encourage "an enormously valuable thaw against the freeze-in of

---

[17] *Therapy*, launching in September 2004,[17] as well as *Arthritis Research & Therapy*, *Breast Cancer Research*, *Critical Care*, *Genome Biology*, and *Journal of Biology*, which are all hosted by BioMed Central <biomedcentral.com>.

[18] www.biomedcentral.com/openaccess/inquiry/myths/?myth=30000

[19] There are many other models, of course. See also David Stern on problems with author-pays models, www.library.yale.edu/science/oa.html .

[20] observer.guardian.co.uk/business/story/0,6903,1168763,00.html. For a recent statement on publishing costs themselves, see the argument by Leo Waaiiers of SURF (Netherlands' higher education and research partnership organisation for network services and information and communications technology, in Nature Sept. 27, 2004: www.nature.com/nature/focus/accessdebate/. Waaiiers estimates that the true, unsubsidized per-article cost to publishers ranges somewhere around $3000 and $4000.

[21] Such as DARE in the Netherlands, www.darenet.nl/en/toon

[22] www.sherpa.ac.uk/romeo.php

[23] www.doaj.org

scientific data." [24] Their goal explicitly goes beyond the sharing of journal articles, to creating more effective communications of all the outputs of research:

> …to recognize, preserve, and extend the historic openness and collaboration that is indispensable to progress in scientific and bio-medical research. In an increasingly proprietary knowledge economy, Science Commons seeks to promote a sharing of information that effectively supports basic research and development of useful innovations, drugs, and other solutions for the public good.[25]

Access and sharing are just the beginning of a new vision of the future productivity of the scientific community. The Science Commons advertisement for a director hints at new directions - investigating the feasibility of an intellectual property rights conservancy, and open standards and practices in bioinformatics research.

Open access opens up much more than access: it enables others to create added value throughout the system – from making articles more dynamic, through post-publication annotation and versioning, to enabling links between data and descriptions, to building the capacity to visualize relationships amongst articles and other resources, to data mining and pattern recognition across sources regardless of where they have been published. Not incidentally, there is emerging evidence that open access leads to greater research impacts across many fields.[26]

In a September 2004 article in *D-Lib Magazine*, Herbert Van de Sompel and his colleagues outline a challenge to move beyond the economic debate over open access, and instead to envision the remarkable opportunities that open access might enable. They call for creating a system of scientific communication that returns the focus from the article or journal, to the scientific process itself.[27] They argue that we are at the early stages of a new technical challenge, of fundamentally reengineering the current scholarly communication system to permit interoperability, support workflow, service sharing, and information modeling. Such a system would be capable of acknowledging datasets, simulations, software, and other dynamic knowledge representations as units of communication, in addition to the journal article. Such a system would also accommodate complex, or compound, documents, made up for example of text, images, and data. They propose that such a system would benefit science by supporting innovation, adaptability, and democratization in the scientific communication system. A key to their proposal, which has attracted NSF funding, is that the current system of communication has become accustomed to losing most evidence of the intimate dynamics of scholarship:

> …The very dynamics of scholarship – the interaction and connection between communication units, authors, readers, quality assessments about communication units, scholarly research areas, etc. – are lost and are extremely hard or impossible to recover after the fact. We feel this loss needs to be remedied in a future … system by natively

---

[24] creativecommons.org/projects/science/proposal

[25] creativecommons.org/projects/science/directorannouncement

[26] Kristin Antelman, "Do open-access articles have a greater research impact?" *College & Research Libraries*, September 2004, pp. 372-382. "The finding is that, across all four disciplines (philosophy, political science, electrical and electronic engineering, and mathematics), freely available articles do have a greater research impact." See also Brody et al., citebase.eprints.org/isi_study/, and Steve Lawrence's research, described at www.nature.com/nature/debates/e-access/Articles/lawrence.html

[27] Herbert Van de Sompel et al., "Rethinking Scholarly Communication: Building the System that Scholars Deserve," *D-Lib Magazine*, September 2004, v.10 n. 9.

> embedding the capability to record and expose such dynamics, relationships, and interactions in the scholarly communications infrastructure.[28]

What a new commons means for the scientific archive: What is striking in this new vision of the future scholarly communication system is that it combines the aims of libraries (access, use) with those of the archives (documenting the context and dynamics of knowledge: people, choices, relationships)[29]. In moving from a document-centric vision of scholarly communication, back to a vision based on the system of science itself, we are faced with a huge range of archival choices. Now that technology provides the capability, we face new choices about what we collect and how we collect it. These decisions in turn will determine the future of the scientific archive: what we'll be able to retrieve, recover, link, and mine, into the distant future. Because the old paper archive model is the best we have, but nevertheless incapable of managing all this brave new content, we have huge challenges to develop the standards of interoperability, format migration, content migration, documentation of re-use rights, not to mention post-production added value in an increasingly living and dynamic archive.

DSpace at MIT began its young life in 1999 as a digital archive for "documents," and this initially meant articles, working papers, technical reports, and other texts. The new face of DSpace research is increasingly turned toward the problems presented by nontraditional research products: databases, data in understood structured formats such as ArcInfo; structured data with emerging standards for interoperability, such as courseware objects; and unstructured data such as images, algorithms, and data models. We have a long way to go! The keys to our strategy include co-development of the DSpace platform among creative partners in the DSpace federation; a perspective that focuses on the people and organizations engaged directly in science, not just publishing; and a commitment not only to more and better ways of seeking and finding scientific research, but a fundamental commitment to the research significance of the archive – to documenting the dynamics and context of science.

3. Opportunities to lose: your stake as chemical scientists

Scientific disciplines, taken broadly, have responded to the era of internet-based communications very differently from each other. They have developed a variety of approaches to scientific communication that uniquely respond to the challenges, responsibilities, and opportunities unique to their discipline. Physics is often said to be the most "progressive" community in terms of scholarly publishing – not only because of its archetypal arXiv preprint server, but because of a history of choices by the APS. Another "progressive" field is said to be computer science. Their model is different from physics – they have a long-standing practice of posting publications on personal web sites. The emergence of Citeseer – also a product of the computer science community has used that practice to create a distributed discovery and access system for their literature second to none in the noncommercial world. A third "progressive" scientific publishing community is biomedicine. Here, the urgency of public interest, of medical need, combined with the vision of the leadership of important funding agencies, have combined with the result that much high profile activity in open access is occurring in

[28] Ibid.

[29] See interview with Cliff Lynch in *RLG DigiNews*, August 15, 2004: "Librarians think in terms of a collection that's a bunch of objects….The goal is to keep the books alive into the future. Archivists look more at aggregate activities and processes that document institutional decision-making. Their concern is less with individual transactions (or) objects and more with looking at whether we have captured…sets of evidence…"

biomedicine (BioMedCentral, PubMed, and the recent position of the NIH on open access for publicly funded research).

Is chemistry not equally progressive? If not, why not? I am told that chemistry is more conservative than other fields, less likely to experiment or change its publishing models. Perhaps that is true.[30]

Chemistry is notably different in several other respects from some other scientific fields. Librarians will tell you from experience that chemists use their own literature – including historic literature – more intensively than other scientists with the possible exception of mathematics. Chemical publications are also the most expensive, on average. One of the reasons for both these factors is that the data contained within chemical publications is a critical source of information for other chemists: whole secondary publishing industries have developed around the job of indexing and extracting data (spectra, properties, etc.) from chemistry journals. The use of special indexes and even the graphical representation of chemical structures are essential to the ability of chemists to grapple with their own documentary record.

Although these features are among those that make chemists (and chemistry librarians) tolerate high prices[31] and value the system that serves them well, these very same features are evidence that chemists stand to gain more than most from new publishing models.

That argument has been made by Peter Murray-Rust and his colleagues, who call for the development of a chemical semantic web, with publications in journals forming a major substrate.[32] If chemical publications were published using standards such as the CML (Chemical Markup Language) family, they would enable the use of computers to accomplish important steps in the scientific process, including establishing data quality, allowing validation, enabling reusability, enabling scale, and ensuring comprehensiveness. Many of these potentialities are stymied now: data are discarded and forgotten, converted to PDF and posted as "supporting information" for an indefinite period; protected on publisher sites by copyright; and accessible or reusable only via

---

[30] "The rift between chemists and biologists may be constitutional. Kornberg describes the biological culture as being dominated by right-brain characteristics and the chemical arena peopled by left-brain folks. "Chemists seem more conservative, analytical and clannish. They focus on molecules: an exotic alkaloid, antibiotic or arcane pigment. They seek the challenge of a molecule with many chiral centers at the very limit of synthetic difficulty and vie to obtain it in the fewest steps with the best yield. They obtain precise data with relatively few and elegant techniques. To them, the chemical monotony of proteins and nucleic acids overrides their biological importance. Biologists on the other hand seem more artistic, eclectic and right-brain dominated. They focus on complex phenomena in cells and organisms, using a wider range of techniques with less precision. They welcome mysteries and complexities and some are disappointed when the veil over a phenomenon lifts to expose molecular detail." Edward J. Walsh, Journal of Chemical Education, 1996 v. 73 n. 12, p. A305, jchemed.chem.wisc.edu/Journal/Issues/1996/Dec/absA305.html

[31] Another reason for tolerating high prices for chemistry journals is said to be the large and well-funded market represented by the chemical industries, including the pharmaceutical industry. Unlike libraries in the academic sector, most such industries have no mission or need to maintain archival collections.

[32] "Representation and use of Chemistry in the Global Electronic Age," Peter Murray-Rust et al, University of Cambridge, www.ch.ic.ac.uk/rzepa/obc

proprietary software. Pointing to examples of machine-readable data in bioscience publishing, these authors encourage chemists to -

> (D)evelop a shared vision whereby information is communal and accessible. It is important to realize that *all information* is potentially valuable and that the producers may not realize at the time what their descendants will require. We argue that the technologies and protocols presented here can be implemented at marginal cost within the publication process, if the community desires. This approach is not novel in other domains. [33]

Specifically, the authors propose that all primary chemical *data* be made available openly at the time of publication, under open access protocols, in a public or institutional repository, with rights to reuse clearly identified with a Creative Commons license. [34]

There is of course another reason why chemists have a strong self-interest to explore new models of publishing and archiving. The multidisciplinary nature of contemporary science, as well as its internationalization, mean that chemists will want to participate as contributors and beneficiaries, to new publishing systems such as those developed by European[35] and Asian colleagues[36], and the publishing systems of the biomedical research community. Interoperability is the key to federating access across emerging national and disciplinary communication systems; efforts such as open access, digital archiving by repositories, and Creative Commons, are the fuel needed to power the value systems of the future.

4. A concluding vision: A Greater Chain of Being:

Both the near- and long-term consequences of the current state of digital publishing of chemistry journals are a threat to the best interests of the scientific community. The best system of scientific communication is not only sustainable (economically as well as technically) but one that best serves the greater purpose of the scientific communication, by sustaining and supporting the ability of the community to both add to and extract value from the cumulating record. That means a system that supports such activities as collaboration, data mining, model sharing, data distribution, and interactive peer review. [37] Such capabilities are not wild-eyed visions: there is good evidence outside and inside the field of chemistry that these things can work.[38] The very idea of a "value chain" is

---

[33] Ibid.

[34] These are not radical, or new, proposals. For a broader context, see the recommendations of the NAS published in their 1997 report, "Bits of Power." They include a call to advocate the principle of full and open exchange of scientific data resulting from publicly funded research, and clarify the importance of sustaining such exchange to the nation's future whenever these forums consider laws that would apply to exchange of scientific data; and for the incorporation of equivalents of 'fair use' as part of regulating access to databases and storage of scientific data and information. *Bits of Power: Issues in Global Access to Scientific Data*, National Academies Press (1997), www.nap.edu/readingroom/books/BitsOfPower.

[35] See for example www.darenet.nl/en/dare_algemeen/toon#wc

[36] Many open access journals are being published in China, Japan, India, and other growing scientific powerhouses outside the U.S.

[37] Reference: U. Pöschl, Interactive journal concept for improved scientific publishing and quality assurance, *Learned Publishing* v. 17, 105-113, 2004.

[38] Poschl, presentation abstract at CINF, 2004: "Substantial improvement can be achieved by an open access two-stage publication process with interactive peer review and public discussion. It enables rapid

restrictive in present, and future, where the value of the content creates a networked field of activities by scientists to use, modify, mine, reuse, assess, and otherwise interact with the record of their science:

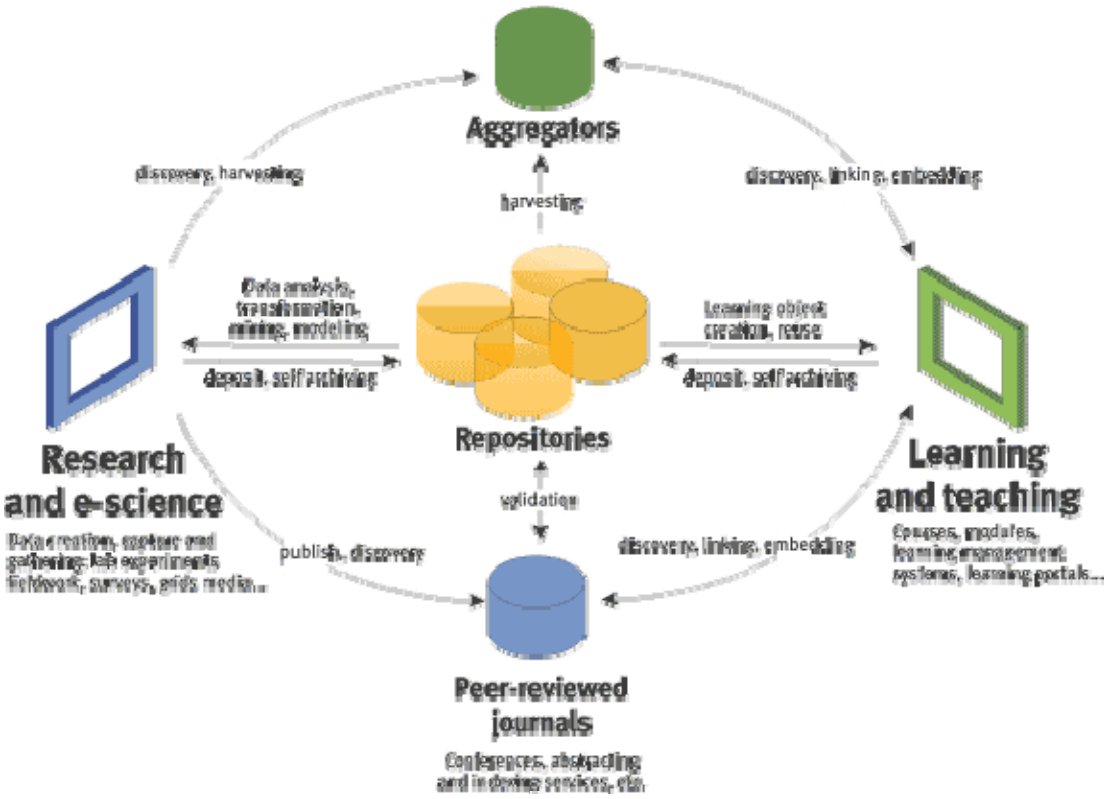

Figure 1: Image from: *OCLC Environmental Scan*, January 2004.

A networked system of communication launched into widespread use just over a decade ago has allowed a return to a vision of scientific communication as a system, not a product. This has profound impacts for you scientists, for publishers, for libraries, and archives. Our roles are all the more interdependent and interconnected. In a recent report by OCLC, Wendy Pradt Lougee is quoted as saying that for libraries, our role as "steward of information goods is being transformed as a collaborator and, potentially, a catalyst within interest-based communities."[39]

---

publication and dissemination of new results in discussion papers followed by thorough and transparent peer review which is open for input from the scientific community (permanently archived and fully citable), and it leads to final revised papers with maximum quality assurance and information density. This approach has been successfully realized and applied in the interactive scientific journal *Atmospheric Chemistry and Physics* (ACP, www.atmos-chem-phys.org), which is edited by a globally distributed network of scientists and has been launched in 2001. The achievements of ACP confirm that the opportunities for interactive peer review and public discussion in a high quality journal are very much appreciated by authors, referees, and the scientific community." See also discussion at: www.researchinformation.info/risepoct04openaccess.html

[39] *2004 Information Format Trends: Content, Not Containers*, OCLC. Cites Wendy Pradt Lougee in footnote 52, "Scholarly Communication & Libraries Unbound: The Opportunity of the Commons," 2004, dlc.dlib.indiana.edu/archive/00001250/

Open access to the literature of chemistry is a key to your ability to redistribute, reuse, add value to, mine, explore, and archive the record of your system of science, without the constraints of paper-based media, commercial ownership, institutional wealth, or fragmentation by publisher.

> At this point in time we can only imagine what is possible, but it is certain that it will dwarf what any one company might achieve. [40]

We will have come a long way when scholarly journals no longer represent products in a make and sell chain of commercial value, but instead stand within a network of interconnected services and activities that support you, the scientists.

Selected reading

Antelman, Kristin. "Do open-access articles have a greater research impact?" *College & Research Libraries*, September 2004, pp. 372-382.

Case, Mary M. "A snapshot in time: ARL Libraries and electronic journal resources," *ARL* August 2004.

Davis, Philip M. "Tragedy of the commons revisited: Librarians, publishers, faculty and the demise of a public resource," *portal: Libraries and the Academy*, October 2003, v. 3 n. 4, pp. 547 ff.

Henry, Geneva. "Online publishing in the 21st century: challenges and opportunities," *D-Lib Magazine*, Oct. 2003, v. 9 n. 10.

Henry, Geneva. Presentation slides on Connexions at *ECDL 2004*, http://cnx.rice.edu/news/News_Item.2004-09-28.1240

Institute for the Future, "Core scholarly information tasks and e-journal features: Expanded discussion," January 2002, prepared for the Stanford University Libraries' e-Journal User Study.

Lougee, Wendy Pradt, "Scholarly communications & libraries unbound: The opportunity of the commons," Workshop on Scholarly Communications as an Information Commons, Indiana University, March 31-April 2 2004.

OCLC Marketing staff, "2004 Information format trends: Content, not containers."

Roosendaal, Hans E. The value of scientific information. Presentation, *ALPSP International Learned Journal Seminar*, 2003, http://www.alpsp.org/2003ppts/roo040403.ppt.

Van de Sompel, Herbert, et al, "Rethinking scholarly communication: Building the system that scholars deserve," *D-Lib Magazine*, September 2004, v.10 n. 9.

[40] www.biomedcentral.com/openaccess/inquiry/myths/?myth=innovation