



Public Access to Government-Sponsored Research

ALA Annual Conference 2007
Monday, June 25, 2007



James King
James.King@nrl.navy.mil
Ruth H. Hooker Research Library
U.S. Naval Research Laboratory

This session explores the 'comparative advantage' of the Federal government as it relates to the Open Access movement. It discusses the implications and imperatives for federal libraries as well as the potential impact on access to government information.

Research at NRL and other Federal research agencies does not typically result in products so the only valid, quantifiable measure of research is productivity found in publications.

Evolution of Scholarly Communication

- Oral teaching, students of the greats
- Manuscripts to extend voice and life of teaching
- Type from printing presses in 18th century to speed creation and dissemination
- Digital and web to restore wide dissemination



We're in the midst of a major step in the evolution of scholarly communication, but not unlike the other steps in the past. Each step in the evolution:

- Introduced wider dissemination of ideas
- Used a new material that was not as durable as its predecessor
- Started with a convergence of a series of innovations
- Caused creation of new industries and dramatic changes to existing industries

David Levy in his book 'Scrolling Forward' defines documents as talking things. With that definition in mind, we can see how the document has evolved – starting with an oral presentation from a great teacher of the past which was limited to a student-teacher relationship and confined to a physical location and time.

We then see that minutes, notes, and proceedings were written down in the form of manuscripts which extended the teaching beyond the time frame and even life of the author but was prone to error and difficult to duplicate. Aggregation of these manuscripts into Egypt's Alexandria library forced the creation of new processes including a classification system, bibliography, and alphabetizing.

The invention of the printing press and the resultant print revolution has so transformed the landscape that Elisabeth Eisenstein's book 'The Printing Revolution in Early Modern Europe' has a difficult time describing what the world was like. Over the past several hundred years, this revolution has created new concepts such as copyright, plagiarism, and editions and has left us with a 'serials crisis' – too many documents for the system to manage.

At the end of World War II, the great scientific research tool that was created to win the war was now turning to peacetime objectives. Vannevar Bush's article "As we may think" helped envision that future and proposed the MEMEX – a microfilm-based desk that contained the entire contents of the Library of Congress – all available at the fingertip. This is coming closer to reality with the digital revolution. Financial issues still remain which is a big reason behind the 'Open Access movement'...

Information Wants to Be Free

“On the one hand information wants to be expensive, because it's so valuable. The right information in the right place just changes your life. On the other hand, information wants to be free, because the cost of getting it out is getting lower and lower all the time. So you have these two fighting against each other.”

Stewart Brand at the first Hackers' Conference in 1984



This is a movement based primarily out of universities that removes price and access barriers to information by capturing pre-prints prior to copyright. Part of motivation of 'movement' is wresting control out of the hands of publishers and back into the hands of scholars.

Open Access/Institutional Repositories

- **Goal:** To provide free access to all scholarly literature to enhance research
- **Journal Focus:** Changing journals to free public access through author fees or 'brief' embargo windows (**Gold**)
- **Institution Focus:** Capturing all pre-published literature created by an institution (**Green**)



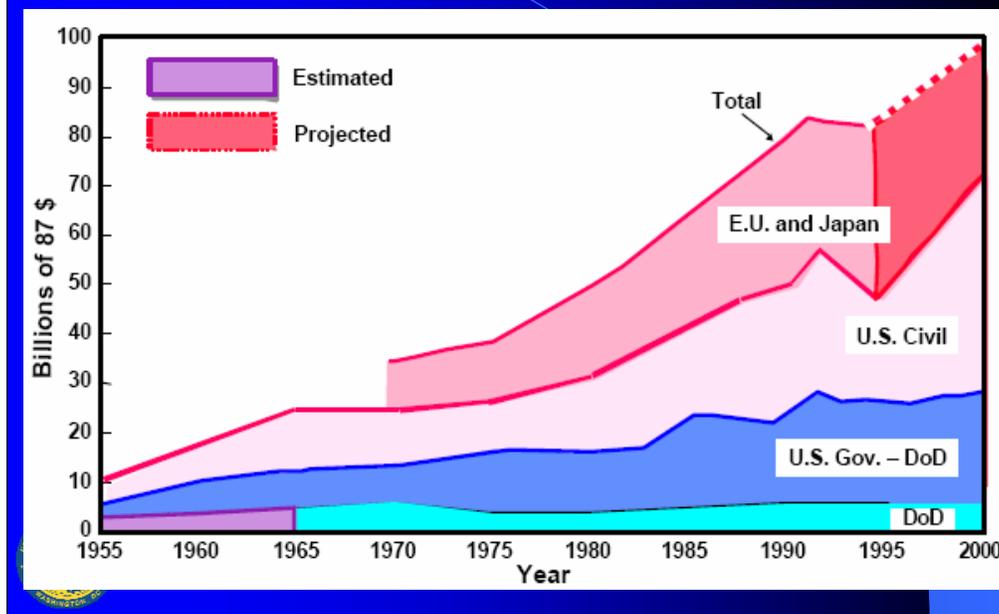
Two traditional approaches as defined in the literature: Gold journals and Institutional Repositories

GOLD: Author fees can range from \$1,500 up to 5 figures per article – with many of these 'publishers' offering an option to purchase an institutional submission license for a fixed fee (like a subscription). Some are concerned that the fees are too low to be sustained when the grant money dries up and that it could lead to lower quality through vanity press. Not easy to find open access articles.

GREEN: An institutional focus, as currently defined and implemented in systems like dSpace, asks authors to deposit final draft of articles prior to publication into a local or domain-specific repository. One goal of this approach is to pull this critical information into a system that will preserve it and broaden access. These systems are often based upon Digital Library standards such as OAI, OAIS, DOI and OpenURL and have several open source and commercial options. Roughly 80% of journal publishers allow authors to do this.

Government has the strategic opportunity and comparative advantage to create a new class or approach – which I propose to call the 'Blue' Open Access method...

'Blue' Open Access



This new class, similar to the GREEN Institutional Repository model, focuses on Government funding agencies. Rather than trying to respond to the scholarly communications crisis with globally-free publications, Blue Access would support Steve Harnad's proposed "Immediate-Deposit/Optional-Access" by ensuring that all government works are (1) captured and preserved, (2) made available to the government research community, (3) provide the raw material to assist with 'research portfolio management' that is growing in Federal R&D, and (4) make portions available to the general public as per copyright regulations.

According to the 2006 NSF Science and Engineering Indicators, the US Federal Government spent \$106.5B on R&D in FY05 – this is more than the top 20 US public corporations (including automotive, pharmaceutical, and IT) combined. In addition, as the chart shows, US Federal R&D spending has been a significant portion of the world's R&D spending, though it has not kept pace with growth over the past decade. These figures point to a significant pool of potential publications that can be captured and preserved through this model. This approach is also retrospective in nature and requires no 'buy in' from the author - each agency can simply identify authors or author affiliation and secure copies of those publications – as far back in time as their able to identify.

As a percentage of the whole, Open Access participation has been very low up to this point - NIH's voluntary submission program has only a 4% submission rate. I believe a major reason for this is fear of having a publication made public before it has been published in the respected publication.

Legislative efforts including the Sabo bill and the Federal Research Public Access Act have attempted to broaden access to publications resulting from federal funding – this approach requires no law changes and can be implemented now.

Chart is from the 2000 Defense Science Board report 'Technology Capabilities of Non-DoD providers' which source data from NSF and the Organization for Economic Cooperation and Development (OECD).

Government & Copyright

- **Government Works**
 - “A work prepared by an officer or employee of the United States Government as part of that person's official duties.” (17 USC 105)
- **Copyright status**
 - Federal Employees – no copyright for official works (17 U.S.C §101 & 17 U.S.C §105)
 - Government Contractors – terms of contract (CDRL)
 - Collaborative Works – murky (17 U.S.C §201)
 - Can't copyright layout/editing (Bender vs. West)
- **Government Purpose**
 - Limited reuse rights over government produced works



<http://cendi.dtic.mil/publications/04-8copyright.html>

Since Copyright law does not apply to government works, Federal agencies have a 'Comparative Advantage' in this effort.

Making the compiled works of feds and contractors available to government agencies will ensure that they are preserved and available to the government research community, will allow management to evaluate the output of basic research – publications, and make portions of the collected works available to the general public.

Efforts like PubMed will naturally be expanded to cover all works sponsored by NIH funding. Other agencies will need to stand up similar efforts or utilize distribution agencies like NTIS or DTIC to aggregate and distribute this information but I believe the critical gathering needs to be done at the local library level to be successful.



The Navy and Marine Corps' Corporate Research Facility

Main Campus - Washington, D.C.

131-acre Campus 111 Buildings

Subject Focus

Physics, Chemistry, Electronics, Oceanography,
Meteorology, Space Sciences

Remote Campuses

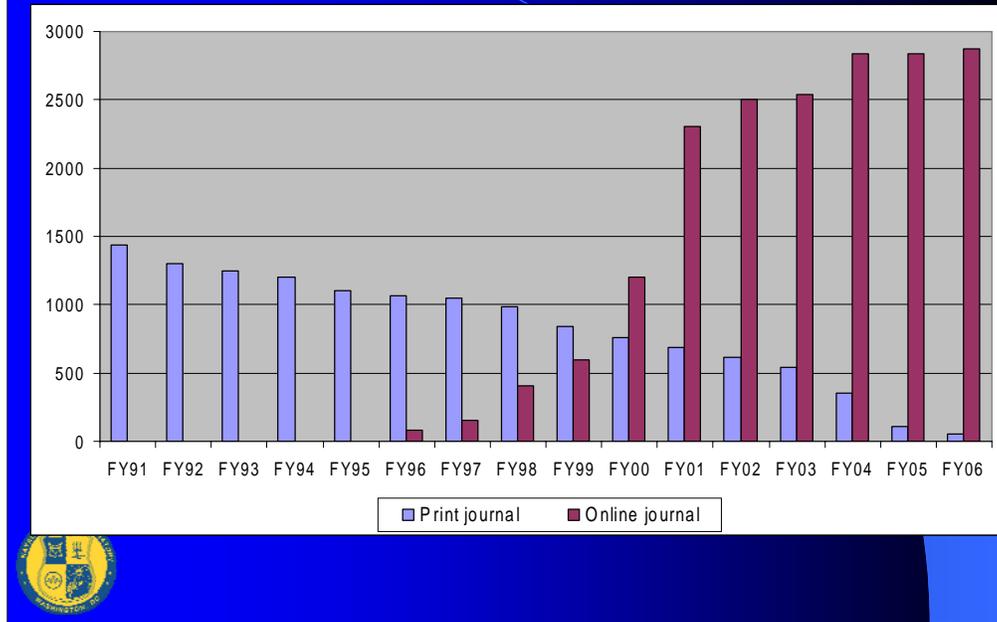
Stennis, MS; Monterey, CA; and Office of Naval
Research



NRL is the corporate research lab for NRL/ONR – Industrially funded (means we get funding from other government funding agencies, including the Office of Naval Research)

NRL's mission is to prevent surprise on battlefield and ensure overwhelming dominance. NRL invented radar, GPS and the first spy satellite.

Journal Collection Evolution



The Ruth H. Hooker Research Library, celebrating its 80th anniversary, has been delivering information to researchers' desktops at NRL & ONR long before the web (1992).

The NRL Research Library has changed the way scientists and engineers do research by providing such a wealth of information at the desktop through useful tools.

For example, 15 years ago, we had 1,400 print journals – we now provide desktop access to double that with less than 2 dozen print titles. 99% of current journals are now provided online-only, 60% of the historical journal collection also online

Most reference sources moving online, resulting in a 100-fold increase in literature and citation searches being done each year compared to the early 90's.

TORPEDO *Ultra* Digital Archive

- 3,000+ journals from 19 publishers
- 6,000 tech reports
- 8M+ articles locally loaded
- Full content search across all content
- Integrated citation linking
- Convera RetrievalWare Concept Searching



TORPEDO Ultra is our locally-hosted digital archive that has been expanded since 1994 when it was built as a research project to see if scientists would accept desktop delivery of journal articles.

This provides NRL with the assurance of permanency since the actual content is kept under local control. It also aggregates all loaded content under one interface, provides more advanced searching than available elsewhere, and is a step towards true digital preservation.

The NRL Library obtains content for TORPEDO in three ways: digitizing agency content, licensing publisher content, and digitizing publisher content. **Digitizing agency content** is the easiest to start with since content created by government agencies is in the public domain, eliminating copyright concerns. **Licensing and/or digitizing publisher content** is more complicated since it requires negotiations with the copyright holders – publishers. This has been done with many of the scientific publishers and societies that we're most interested in.

Capturing corporate knowledge

NRL Publications Database v3.0β
A collection of publications produced by researchers of the United States Naval Research Laboratory.

Total Items **54,896** Journals **30,089**; Books **206**; Book Chapters **282**; Conference Proceedings **5,530**; Abstracts **2,909**; Patents **2,140**; Technical Reports **13,738**; Invited Talks **1**; Keynote Presentations **1**;

SEARCH Any Author Search

phrase all word(s) any word exact First Author
Author Name Format: **Lastname, Initial. Initial.**

▶ **REPORTS / STATISTICS**
• Patents Citing NRL Papers
• View the 25 most cited papers by Web of Science

▶ **ADD PUBLICATION** - Paste or type citation data to add a missing publication.

NRL PUBLICATIONS BY YEAR browse by year (ex. 1975) or range **GO**

2000 ·1733·	2001 ·1656·	2002 ·1739·	2003 ·1589·	2004 ·1519·	2005 ·1295·	2006 ·1191·	2007 ·229·		
1990 ·1298·	1991 ·1324·	1992 ·1342·	1993 ·1672·	1994 ·1849·	1995 ·1683·	1996 ·1726·	1997 ·1932·	1998 ·2019·	1999 ·1820·
								1988 ·1087·	1989 ·1039·

After WWII, most critical government research was captured in research reports and distributed through places like DTIC and NTIS. Now, critical information is residing in corporate databases, MS Word files, PowerPoint slides, etc. that must be retained but the volume negates using traditional library processes. Deanna Marcum, Assoc. Librarian of Congress recently questioned the future of established cataloging. Trend with IR has been to create a new 'silo' of unpublished information rather than folding into a continuum of scholarly information - a 'formalization of the informal'.

The NRL Research Library's goal is to:

- capture the 'intellectual output' of NRL from beginning (1923)
- Capturing already published info – no author intervention required + historical info
- Currently contains traditional NRL publications (journal articles, conference proceedings, and book chapters)
- Adding support for issued patents, technical reports, laboratory notebooks, oral presentations, invited presentations
- Changing 'chop chain' to capture thru required processes
- Also working w/AO to capture lists @ higher group level w/goal of reducing work of AO
- Providing easy web tool for reviews & promotions (annual summary)
- Bragging rights of Top 25 most cited publications
- Expertise list by reviewing pubs and citation counts – can search keywords or TORPEDO
- Creating compendium of pubs for retiring scientists before leaving
- Scanning lab notebooks and interviews of prominent scientists
- Moving towards mandating submission in order to use with annual reviews and promotions
- Digital preservation
- Virtual journal of Naval research

Making system indispensable for scientists & administrators – ensuring that a portion of the corporate memory is retained for the future.



This is the Chinese word for 'CRISIS'

The first symbol represents 'Danger' – libraries are facing a lot of them:

- Baby Boomer retirements
- Budget cuts/downsizing
- Lack of corporate knowledge retention
- Old school librarians focused on preservation of the status quo

The second symbol is 'Opportunity' – we have several of them ahead of us, depending upon our perspective and response to the changing environment:

- Lots of retirements = opportunities to hire new people with new ideas
- Advantage of government works and copyright
- Semantic Web requires authoritative data sources to 'mashup' – the Feds have tons of it!
- Teleworking will REQUIRE a digital library
- Working for the Federal Government allows you to do something that really makes a difference, something of value

"It is time for a new generation of leadership to cope with new problems and new opportunities. For there is a new world to be won." John F. Kennedy