

# Scientific communication: evolution or paradigm shift?

Allan Foster

INFORUM, Prague, 23<sup>rd</sup> May 2006

[allan@allanfoster.co.uk](mailto:allan@allanfoster.co.uk)

[www.allanfoster.co.uk](http://www.allanfoster.co.uk)

# Communications occasionally go awry ...



Source: BBC24 television channel, May 2006

# Structure of presentation

- Strains in the existing pattern of scientific communication
- Data management challenges
- Open access: where are we?
- The new information companies
- Social technologies: are they changing the information world?

# Scholarly communication: 1970s-

- increasing reliance on journals as main channel for dissemination of scientific knowledge, growth paralleling that of research produced
- dominance of 'reader-pay' or 'library pay', as opposed to 'author-pay' model of journal dissemination
- existence of many publishers in market, with two big groups of publishers, For-profits (FP) and Not-for-profits (NFP - including learned societies and university presses)
- very fast growth of some big FP publishers, through new journal introduction, administering journals for learned societies, and through mergers.

# Tensions in traditional pattern of scholarly communications

- Economic & financial change – for buyers, contributors and providers
  - Pressures on universities & funding bodies, incl. selectivity, re-focusing, costing, static library budgets
  - Pressures on ‘researcher as author’
  - Pressures on publishers – costs, markets, margins, technology & voluntarism (editing & reviewing)
- Technological change
- Social change

# Scholarly communications market now

- Core STM publishing: \$7-11B
- 1975-1995 journal prices increased by 300% over inflation rate
- Very imperfect market mechanism
- FP journals are 3 times more expensive than NFP titles
- Positive link between quality of journal and its price

Source: **Study on ... scientific publication markets in Europe**. EC, 2006.

<http://digbig.com/4hrnh>

# EU Survey of ... scientific publication markets in Europe: recommendations

- Guarantee public access to publicly funded research results shortly after publication
- Aim at a level playing field in terms of publishing business models
- Extend quality rankings of scientific jnls
- Guarantee perennial access to scholarly journal digital archives

# EU Survey ... policy recommendations (cont.)

- Foster interoperable tools to improve visibility, accessibility & dissemination
- Promote pro competitive pricing strategies
- Scrutinise future significant mergers
- Promote the development of electronic publications (VAT treatment; public/private partnerships)



# Data management: some key issues

- How can the results of publicly funded scientific research be made accessible, optimally, internationally?
- How can we provide suitable integration and linkage between scientific publication and underlying data sets, incl. observational and experimental data?
- How can we encourage the proper curation of important data to ensure long term accessibility?
- How can we best work with partners in this endeavour, including funders, universities, researchers, users and publishers?

# Data management issues (cont.)

- Humanities and social scientific data as well as STM
- Legal and IPR issues to be faced in these tasks
- Very mixed picture across European states on advice/good practice on ownership of data, restrictions on use...
- Natural interest by funders and researchers to move on to new areas without sufficient attention to the established research record
- Mainly cultural, sociological and financial factors rather than technical
- Highlight data management as a profession and career

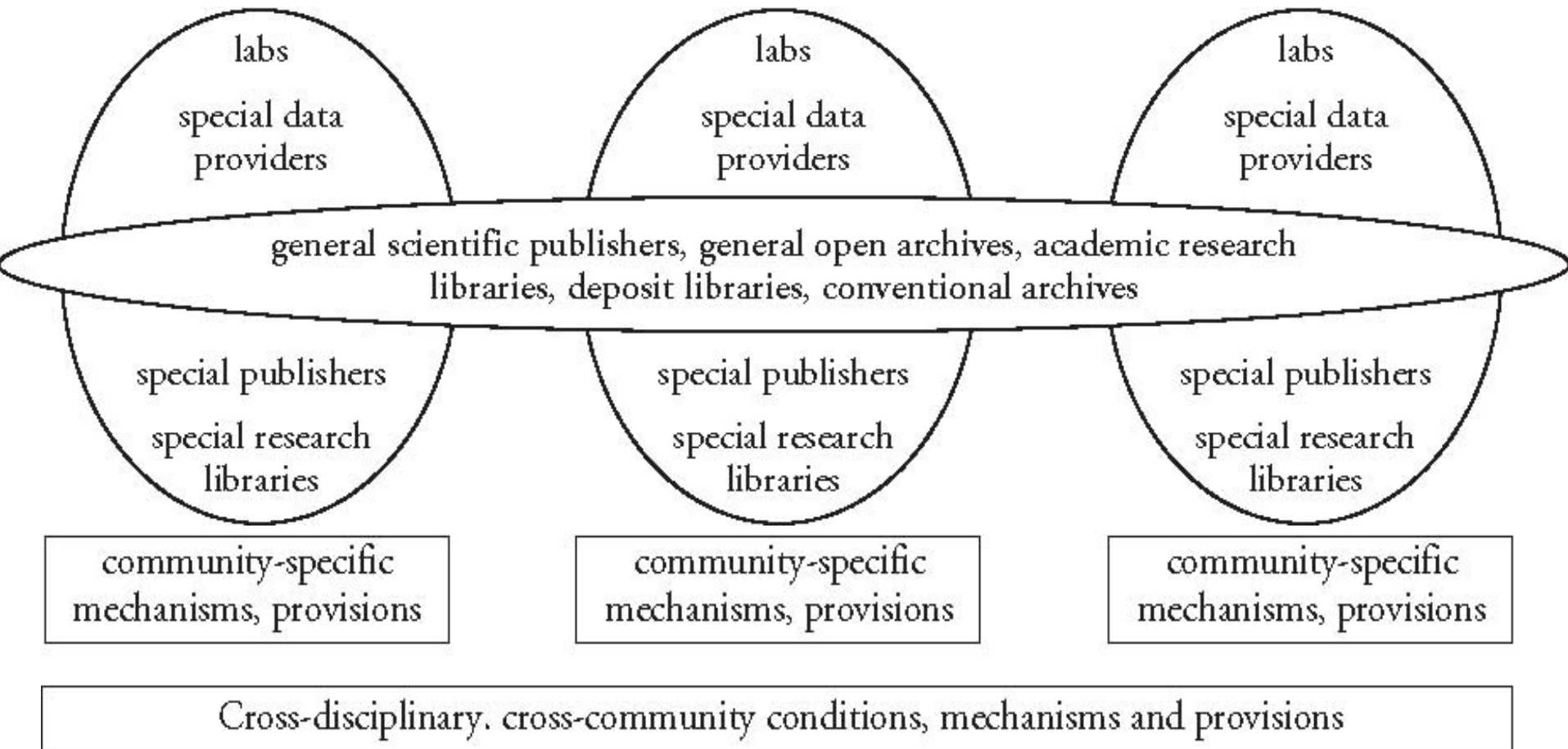
# Data curation

- Key to sustainability, 'reproduce-ability' and re-use of reliable and trusted digital resources
- Incentivising good practice: research funders requirements & career recognition of data management outputs
- Reproduce-ability sometimes requires a curation of a 'fixed' rather than dynamic dataset

# Data curation (cont.)

- Software and versioning issues
- Need for new accounting model recognising long term costs of data curation and preservation
- European and international ongoing work such as by UK's Digital Curation Centre (DCC) to undertake research and disseminate good practice

# Model of virtual infrastructure



Source: Task Force Permanent Access. Permanent access to the record of science: strategic action programme. ETFPA, 2005.

# The Open Access agenda

- Open access journals
- Self-archiving
- Repositories

# BOAI definition of an open access journal

“**free availability** on the public internet, permitting any users to **read, download, copy, distribute, print, search, or link** to the full texts of these articles, **crawl them** for indexing, pass them as **data to software**, or use them for any other lawful purpose, **without financial, legal, or technical** barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give **authors control over the integrity of their work** and the right to be **properly acknowledged and cited.**” (my highlighting)

Budapest Open Access Initiative FAQ: <http://digbig.com/4hrbg>

Directory of open access journals - Microsoft Internet Explorer

File Edit View Favorites Tools Help Links Customize Links Free Hotmail RealPlayer Windows Windows Marketplace Windows Media my delicio.us post to delicio.us

Address http://www.doaj.org/ Go Google gadwin print screen Search PageRank ABC Check AutoLink AutoFill

Back Search Favorites

# DOAJ DIRECTORY OF OPEN ACCESS JOURNALS

[Find journals](#)  
[New titles](#)  
[Search articles](#)  
[Suggest a journal](#)

Welcome to the Directory of Open Access Journals. This service covers free, full text, quality controlled scientific and scholarly journals. We aim to cover all subjects and languages. There are now **2232** journals in the directory. Currently **633** journals are searchable at article level. As of today **97184** articles are included in the DOAJ service.

[About](#)  
[Questions & answers](#)  
[News](#)  
[Cooperation](#)  
[Contact](#)  
[Support DOAJ](#)  
[OpenDOAR](#)

**For journal owners**

**Browse by title**

[A](#) [B](#) [C](#) [D](#) [E](#) [F](#) [G](#) [H](#) [I](#) [J](#) [K](#) [L](#) [M](#) [N](#) [O](#) [P](#) [Q](#) [R](#) [S](#) [T](#) [U](#) [V](#) [W](#) [Y](#) [Z](#)

**Browse by subject**

<a href="#">Agriculture and Food Sciences</a>	<a href="#">Arts and Architecture</a>
<a href="#">Biology and Life Sciences</a>	<a href="#">Business and Economics</a>
<a href="#">Chemistry</a>	<a href="#">Earth and Environmental Sciences</a>
<a href="#">General Works</a>	<a href="#">Health Sciences</a>
<a href="#">History and Archaeology</a>	<a href="#">Languages and Literatures</a>
<a href="#">Law and Political Science</a>	<a href="#">Mathematics and Statistics</a>
<a href="#">Philosophy and Religion</a>	<a href="#">Physics and Astronomy</a>

start 10 Internet E... 4 Firefox Joho the Blog - ... Address Book 2 Microsoft Of... 2 Microsoft Of... EN 08:52

Source: DOAJ, [www.doaj.org](http://www.doaj.org)

Screen shot 18/5/2006



# JISC's open access programme

- Limited funding to allow authors to experiment with OA publishing
- 6 participating publishers: PLoS, Institute of Physics, J. of Experimental Botany, Inter. Jnl of Crystallography, BMJ, OUP
- 124 authors; 78% would publish in an OA journal again. Factors include:
  - Free access for users
  - Wide exposure and consequential academic discourse
  - Prestige
  - Impact factor
  - Citations
- Publishers: performance positive; OA experience fruitful
- Authors 50:50 split on whether fees should be paid by govt agencies or from research grants

Source: Evaluation of the JISC's open access funding initiative. Key Perspectives Ltd, Spring 2006. See <http://digbig.com/4hrfa>

# Repositories

- Research publications, data and learning/teaching resources
- Institutional *and* disciplinary
- Inter-relationships between institutional repositories and specialist data centres
- “Scientific data is best handled by scientists in data centres” (NERC - UK)
- Open access embraces self-archiving in repositories *and* new business models for publishers

# Repositories (cont.)

- Currently, how interested are researchers and institutions in repositories??? (Some hostility)
- Should deposit be a requirement by funders? (NIH, RCUK & Wellcome policy directions)
- Once again, cultural change is essential
- **D**igital **R**epository **I**nfrastructure **V**ision for **E**uropean **R**esearch (DRIVER)
  - Initially a common network of 51 existing repositories in 5 countries
  - Collective and enabling service layers
  - Using standards such as OAI-PMH, persistent identifiers and some technology standards (SOA, web services)

# The new information behemoths

- Google
  - Google Scholar
  - Google Books
  - Google Co-op
  - Google Trends
  - AdSense: matching adverts to keywords on e-journal page. Payment by click-throughs
- Yahoo!
  - with Microsoft, RLG etc (Open Content Alliance): *Open Library* – cataloguing & digitisation
  - Social technologies – Flickr (photos), Del.icio.us (tags)

# Google's superior searching

via a multi-algorithmic approach,  
recognising value of:

- **Full text:** matching keywords in the query and content
- **Information architecture:** analysing the internal link structure and hand-crafted metadata of each web site
- **Free tagging:** using the links between web sites

# Google Trends: comparing searches

Google Trends: open access, repositories - Microsoft Internet Explorer

Address: <http://www.google.com/trends?q=open+access&>

Search: open access, repositories

Tip: You can compare searches by separating with commas.

### Trend history

All Regions All years

● open access ● repositories

Search volume

News reference volume

2004 2005 2006

- A** [BioMed Central explodes Open Access myths](#)  
EurekAlert (press release) - Mar 22 2004
- B** [Report Seeks Open Access Research Policy](#)  
Wilkes Barre Times-Leader - Sep 10 2004
- C** [Open access moves a step closer](#)  
EducationGuardian.co.uk - Feb 2 2005
- D** [Academics thrash out open access details](#)  
EducationGuardian.co.uk - Mar 3 2005
- E** [Open access online veterinary journal launches](#)  
EducationGuardian.co.uk - Jun 1 2005
- F** [New study weighs impact of open access on scholarly journals](#)  
EurekAlert (press release) - Oct 11 2005

Cities Regions Languages

Top cities (normalized)

1. <b>Bangalore</b> India	
2. <b>Adelaide</b> Australia	
3. <b>Washington</b> United States	

Internet

start 11 Internet E... 4 Firefox Joho the Blog - ... 2 Microsoft Of... 2 Microsoft Of... Microsoft Excel -...

EN 09:55

# Social technologies

- Weblogs
- Tagging
- Wikis – public & private
- RSS & syndication

# Weblogs

- as a marketplace of ideas
- as scholarship
- as democracy
- as a challenge to traditional authority
- for news
- as a corporate communications tool
- for understanding corporate reputation



# Are information professionals taking blogs sufficiently seriously?

“Rethinking what is a ‘local collection’ is very necessary in an age in which less and less content starts out in book and journal format and more in raw forms such as weblogs.

In the meantime a growing army of webloggers are performing this function to some degree already themselves, cataloguing the Web from very personal and oftentimes local perspectives.

These people are creating their own ‘sweet science’ of content organization that's influencing more and more readers and writers worldwide.”

# Peer review v social collaboration

- Are collaborative technologies antithetical to peer review? Can they co-exist when blogs challenge traditional notions of authority?
- Connotea from Nature Publishing: designed for scientists <http://www.connotea.org>

# Connotea: organising, sharing & discovery service for scientists from Nature Publishing Group

The screenshot shows the Connotea website interface within a Microsoft Internet Explorer browser window. The browser's address bar displays `http://www.connotea.org/`. The website header features the Connotea logo on the left, a search bar with a dropdown menu set to 'All', and a 'Search' button. To the right of the search bar, it indicates the user is 'logged in as AllanFoster' with a 'My Library' link and a 'Logout' button. Below the header is a navigation menu with links for 'Home', 'Latest news', 'About this site', 'Site guide', 'FAQ', and 'Contact us'. A secondary menu highlights 'Popular links' and 'Recent activity'. The main content area is divided into three sections: 'Organize your references' (with an image of a binder), 'Share with others' (with an image of hands holding a globe), and 'Discover new leads' (with an image of a person's legs in a boat). A central text box contains the following information: 'Connotea is a free website to help researchers and clinicians manage and share information. Watch a [short video](#) about Connotea. [Register](#) to get started.' The 'nature.com' logo is visible in the bottom right corner of this text box. The Windows taskbar at the bottom shows the 'start' button, several open applications including '10 Internet E...', '4 Firefox', 'Joho the Blog - ...', '2 Microsoft Of...', 'Inforum 1', and 'e-IRG\_data\_ma...', and the system tray with the time '11:04'.

# What do we know about HE students' use of discovery services? [INFORUM 2002]

- heavy reliance on generic search engines - the 'Google Factor'!
- want fulltext
- extremely 'quality tolerant'
- simple search strategies
- perplexing wealth of quality (paid-for) web resources, let alone free Internet stuff
- 'satisficing' behaviour ie 'good enough'
- and many academic staff behave in the same way!

**All pretty much true in 2006**

# Searching and taxonomies

Librarians' views of researchers' use of discovery tools:

- variable and idiosyncratic
- past experience constrains choice
- untroubled by different interfaces
- free text searching is norm except in a few areas such as chemistry
- little use of formal thesauri
- the Google inheritance - or has it always been so?
- some major exceptions eg systematic reviews

# Searching and taxonomies

- 'Folksonomies': collaborative subject analysis
  - Tagging
  - Tag clouds
- Collabulary – a compromise between top down taxonomies and casual tagging
- A world of messiness – piles of leaves, not trees (David Weinberger)  
(see <http://www.hyperorg.com/blogger>)

# Technorati's home page tag display: 19<sup>th</sup> May 2006

The screenshot shows the Technorati website in a Microsoft Internet Explorer browser window. The address bar displays <http://www.technorati.com/tag/>. The page header indicates "Currently tracking 40.1 million sites and 2.4 billion links." and includes navigation links for "Member Sign In", "Sign Up", "Help", and "About".

The main navigation bar features the Technorati logo and four tabs: "Search", "Tags", "Blog Finder", and "Explore". Below this is a search input field with a "Go to Tag" button and an "Options" dropdown.

The left sidebar is titled "Today's Most Popular" and is divided into three sections:

- News**
  - Forever Pregnant
  - Forever Pregnant
  - USATODAY.com - NSA has ma...
- Movies** (with a small image of a movie poster)
  - The Da Vinci Code
  - Mission
  - Poseidon (2006)
- Books** (with a small image of a book cover)
  - The Da Vinci Code
  - The Pi-Calculus
  - The Da Vinci Code

Below these sections is a "Top 100 Blogs" list:

- 老徐 徐静蕾 新浪BLOG
- Boing Boing
- Engadget

A "More in Popular >>" link is located at the bottom of the sidebar.

The main content area features a large heading: "Tags: The real-time web, organized by you". Below this is a paragraph explaining the concept of tags: "A tag is like a subject or category. This page shows the most popular tags, starting with the ones that are hottest today. The bigger the text, the more active that tag is. We're currently tracking 6.9 million tags. Do you blog? Find out how to **tag your posts**."

A "Hot Tags This Hour" section is highlighted in a light green box, displaying a list of tags with varying font sizes and colors to indicate their popularity:

Advertising ... American Idol ... Baseball ... **Brazil** ... Bush ... China ... Da Vinci Code ... **Erectile Dysfunction** ... Flickr ... **Health News** ... **Immigration** ... Iraq ... macbook ... **men\'s health** ... Microsoft ... MLB ... non-classe ... Sex ... **Sexual Health** ... thursday thirteen ... **Viagra** ... web-20 ... web2.0 ... wordpress ... Yahoo ...

Below the hot tags is a section for "Top 100 Tags from A to Z(All Languages)", which includes a list of tags such as "Allgemein ... Apple ... Art ... Art and Photography ... Articles ... **Blog** ... **Blogging** ... Blogs ... book ... **books** ... **Business** ... Computers and ...".

The bottom of the browser window shows the Windows taskbar with the Start button, several open applications (Internet Explorer, Firefox, and others), and the system tray displaying the time as 10:46.

# Factiva's use of tag displays

Factiva | Beta Manage Alerts | Feedback | Help

Search  Search Search Within Filters

Filters: **Company: John Lewis Partnership PLC** **Company: Marks & Spencer Group PLC**

Did you mean: **Marks & Spencer Group PLC**

**Publications** Alert me **Discovery Pane**

All | [Newspapers](#) | [Magazines](#) | [Newswires](#)

[MKS.L](#) Marks & Spencer Group PLC **519** ▲.25 LSE GBp [Chart](#) | [Snapshot](#)

Headlines 1 - 20 of 59 | [Next](#) ▶ Sort by: **Relevance** | [Date](#)

- Marks & Spencer sparks optimism among retailers**  
Scotland on Sunday, Jan 8 2006, 821 words, (English)  
...MARKS & Spencer is expected to provide some cheer for high street retailers this week by confirming that its recovery remains on-track following improved trading over Christmas.  
...
- Investors Chronicle: Christmas crackers and trading turkeys revealed.**  
Investors Chronicle, Jan 13 2006, 1330 words, (English)  
...Mid-market outlets Marks and Spencer (M&S) and the unlisted John Lewis Partnership reported like-for-like sales increases of 2.9 per cent and 7.1 per cent, respectively, on Christmas last year. At Majestic Wine, which sells cases of champagne...
- Women: CORKING! AND LESS THAN £5**  
Mirror, Feb 7 2006, 490 words, (English)  
...Marks & Spencer Pinot Noir 2005 Casablanca Valley ...
- Heinz, M&S get funds to reduce household waste**  
The Grocer, Jan 14 2006, 190 words, (English)  
...Heinz and Marks and Spencer have both received a chunk of the latest funding from the Waste & Resources Action Programme (Wrap) to reduce household waste. ...

**Date**

Distribution: Weekly

**News Clusters**

chicken sandwiches : breast meat : chicken sandwich reduce household waste : ready meal chicken : glucose syrup : added ingredients uk consumer : packed sandwiches : water starch profit warning : sales growth : weeks

**Companies**

John Lewis Partne...	59
Marks & Spencer G...	59
J Sainsbury PLC	8
Dairv Crest Group ...	3



# Summary: some key points

- Academics (and research funders) want more control over their IP (institutional & personal) yet still seek scientific status & recognition through high impact journal publishing
- Data management: more link ups between scholarly publication and original data; greater professionalism in data curation
- Richer & more varied business/publishing models, embracing high quality publishers
- New, powerful entrants to scholarly content access & distribution making their mark

# Summary (cont.)

- Open access here to stay but with turbulence
- Rapid, uneven growth for OA journals & repositories (incl. links to scientific datasets) – and interoperable standards
- Social technologies are complementary to quality peer reviewed sources
- Taxonomies, ‘folksonomies’ and ‘messiness’ (complexity)

Thanks for listening!

Allan Foster

[allan@allanfoster.co.uk](mailto:allan@allanfoster.co.uk)

[www.allanfoster.co.uk](http://www.allanfoster.co.uk)