Information Services & Use 27 (2007) 3–34
IOS Press

Understanding user behaviour and its metrics

The needs of scholarly authors and researchers
ICSTI Winter Public Conference
The British Library Conference Centre
January 22nd 2007

David Brown [a] and Bernard Dumouchel [b]
[a] *Head of Scholarly Communications and Innovation Support, The British Library*
[b] *Former Director-General, Canada Institute for Scientific and Technical Information*

Welcome

Lynne Brindley, Chief Executive, the British Library

We are pleased to bring together such a great group of people to speak about various aspects of user behaviour and trends, an important topic for us all. We are sometimes guilty of assuming we know what users want. Those assumptions have been challenged in recent years by sophisticated evidence-based studies, the impact of which has been further heightened by digital systems that are responding to the new and changing needs of researchers operating on the front lines of science. To start the programme, two such scientists will provide accounts of how they respond to the information challenges they face in their respective fields. During the day we will learn about the background motivations for, and situations in which, researchers increasingly operate. The challenge is to interpret what this means in terms of our own organisations. This is a timely issue for the British Library as it revisits its science, technology and medical strategy to meet its own changing needs as well as the needs of new generations of researchers. This will require adaptations to our information and knowledge services and policies, something I suspect applies to all ICSTI members here engaged in service provision. A relevant example of this is the British Library's recent launch of the UK PubMed Central, a US–UK collaboration with the support of research funding agencies and driven by the Wellcome Trust. We are pleased that representatives of those organisations are here today. Participants from 12 countries are among us, making this a truly international gathering. I am delighted to welcome you all here today.

Introduction

Bernard Dumouchel, Director-General, Canada Institute for Scientific and Technical Information (CISTI/CNR Canada)

I want to express our appreciation to the British Library for having us here today and to thank David Brown of the British Library for organising this conference. We have a number of impressive speakers, each of whom brings significant experience to the topic.

ICSTI – the International Council for Scientific and Technical Information – offers a unique forum for interaction between organisations that create, disseminate, and use scientific and technical information. It is a truly international organisation and it is a pleasure to welcome participants from so many different countries. As part of its approach in delivering its mission, ICSTI organises two public conferences a year (in spring and winter) where issues are discussed in a public forum. If you visit our website you will see ICSTI's organisational activities. This year ICSTI is focussing on assessing the quality and impact of research. Our next public conference will be held in Nancy, France on June 21, 2007, the focus of which will be practices and trends in information as they relate to user behaviour and impacts.

Today we will look at the challenges raised by the emerging paradigm of information overload and new methods of assessing and using STM information in relation to the user. The first part of the conference will focus on user behaviour *per se* and how information services must adapt to emerging trends. We must attend to what our users are requiring of us and consider how can we adapt our services and approach in SciTech and medical information. Subsequently we will discuss the relevance of information services as measured through metrics. Finally, we will consider funding as a driver in influencing user behaviour and the impact of funding agencies.

SESSION I
USER BEHAVIOUR

Identifying, accessing and disseminating scientific information: a biological scientist's perspective

Dr Sarah Coulthurst, Cambridge University

This presentation will provide a biological scientist's perspective on identifying, accessing and disseminating required scientific data. The information requirements of a researcher vary from day to day depending on personal preference, field, types of techniques used, career stage and project cycle. Despite this individuality, there are some common problems and approaches.

Accessing information

The biological scientist seeks two categories of information: targeted/specific and speculative. The most important method for locating information is citation database searches (PubMed, Web of Science) and search engines (Google Scholar). Some drawbacks of keyword searching are two or more things with the same name, common author surnames, lack of specificity of search terms, and multiple search item names. Other search methods include manual examination of tables of contents (TOCs) of key journals, automatic updates (electronic TOC alerts, Pub Crawler), traditional citations in related articles, research highlights in "news and views", articles in other journals, and word of mouth (at conferences, seminars, journal clubs).

Primary source information acquisition via international peer reviewed journal articles is increasingly important in certain areas and there is growing interest in supplementary journal article information (e.g. moving images or detailed raw data) that may not be appropriate for inclusion in the main article but is often available from the journal or at the author's website. Some concerns with supplementary information include poor formatting and a tendency among high profile journals to relegate key information to the supplementary section due to space constraints.

Publicly-available scientific databases (e.g. genome sequences) are a key source of data that is appropriate for inclusion in journal articles. However, appropriate depositions of data and proper citation of articles in databases is key to quality maintenance. When a particular database entry is associated with a quality article, we know that expert peers have examined the data and this serves as a form of validation.

Journal selection

The following criteria affect journal selection: the journal's area of interest, reputation/profile, impact factors/citation indices, funding agency requirements, how much the journal charges for open access, page charges, style and layout, availability and previous experience. Our perception of the publishing experience as positive or negative is based on the clarity of instructions to authors, the extent to which contributors are kept informed of progress, speed and fairness of refereeing and editorial decisions, accuracy of final article and quality of communication between author and journal.

Journals and journal websites now offer a broad range of services that include news and opinion articles, debate fora, job boards and funding opportunity announcements, conference and course announcements, technical updates and educational tools. All of these raise the profile of the journal within the scientific community.

Conclusions

Peer review in high quality international scientific journals is fundamental in acquiring and disseminating trustworthy and high quality information. Multiple routes exist to identifying papers of interest, and information acquisition varies across people and over time. Articles are usually obtained electronically and the occurrence of accompanying supplementary information is increasing. Proper interaction of journals with well-maintained public databases is crucial for ensuring the high standards of both. Multiple factors are involved in the selection of a journal for publication, and good communication with the journal is key to a positive experience. The scope and profile of a good journal extends beyond the publication of scientific papers.

Examples of geological research: views and anecdotes

Dr Adrian Jones, University College London

This presentation is from the perspective of geological sciences and will provide a snapshot of two levels of research interest: short term (project conducted by MSci students) and long term (programme conducted by PhD/postdoctoral students). I will consider the example of volcanic CO_2 emissions, a topical area that demands interdisciplinary and cross-sectoral interaction.

Volcanic CO_2 can be measured and one can see how it is dispersed. The scale of an eruption and the way in which CO_2 is released depends on the scale and nature of the eruption, its duration, concentration and source. Not all volcanoes produce the same amount of CO_2, but virtually all volcanoes produce CO_2 in some quantity. Volcanic activity takes place at two paces – the steady state (which is what we see today – we are currently in a volcanic low) and catastrophic (includes volcanic events that have overwhelmed the planet and contributed to mass extinction).

The two projects I will discuss in the context of data collection both have immediate interest. We can examine rocks in the UK that are part of a cycle of activity connected to a significant global problem. Figure 1 shows the opening of the North Atlantic while the map on the bottom right shows the geographic

Fig. 1. About 60 Ma ago: The North Atlantic *Large Igneous Province* what started it?

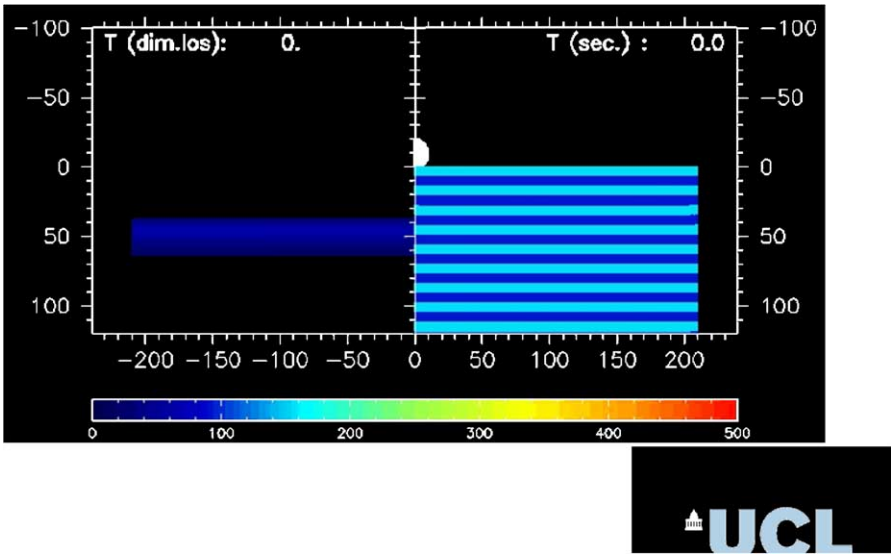


Fig. 2. Could giant meteorite impacts be the triggers for catastrophic volcanism (and mass extinction events)?

situation as it would have been prior to the existence of the North Atlantic some 60 million years ago. The gradual westward movement of Greenland created a gap that is filled with what today is known as the North Atlantic Ocean (underlain by basaltic lava).

A catastrophic event, such as a large meteorite impact, could have generated large amounts of melting on the earth that overloaded the climate with material, including CO_2. In the past, large projectiles, some of them 20 km in diameter, collided with the earth. Figure 2 illustrates that the impact or "splash" made is 100 km in scale – a gigantic event equivalent in size to that which is thought to have led to the dinosaurs'

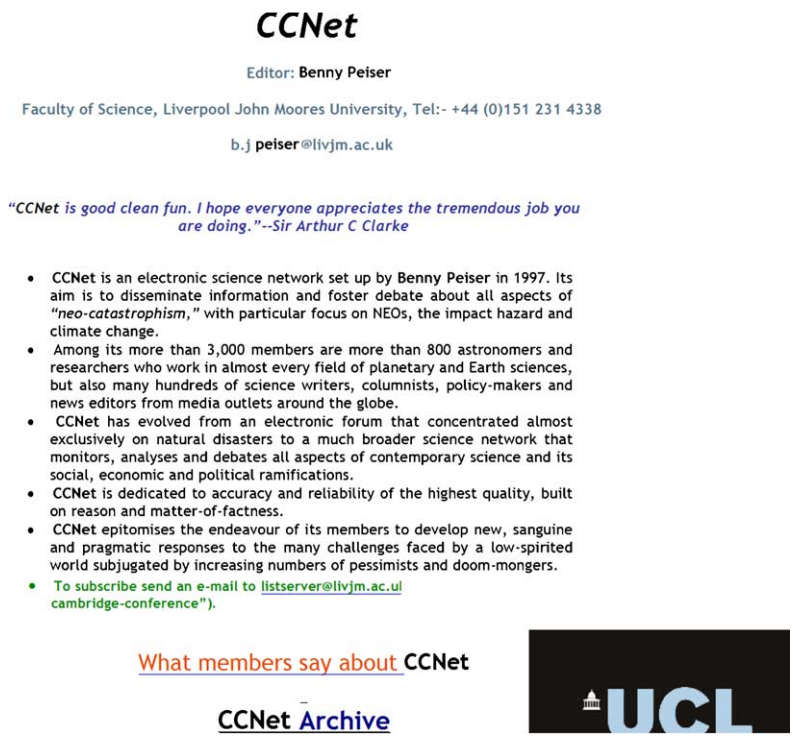


Fig. 3.

extinction. This is not an unreasonable size and we think we can pose this as a mechanism to explain large bodies of magma which existed on the Ontong Java plateau 120 million years ago – twice the age of the 60 million year impact. Million or Billion?

CCNet (Fig. 3) has grown from the monitoring of catastrophic events to the analysis of their meaning and provides a useful tool for noting what is taking place with respect to newsworthy events, scientific publications and government monitoring.

Figure 4 shows the first page of a hit list of a student enquiry for a two-term project looking at rocks involved in the North Atlantic. Students collected lava samples at the base of the Isle of Mull. Google can be effective in identifying topical information – even in a simple-minded search.

I want you to reflect on some key aspects of geological research. Published geological data can offer vital accounts (which allow for the scientific treatment) of rocks which have vanished or events for which no other records exist. 3D and 4D problems are also helpful and the planetary dimension is vital. We are involved with high science technology and the development of new analytical equipment. Geological research often requires making decisions or models from incomplete data. Other search tools in regular use include ScienceDirect, Web of Knowledge and Web of Science.

We routinely obtain data from the societies, research councils and partner institutes (e.g. Natural History Museum). We also call upon large-scale facilities such as those relating to diamonds, synchrotron, ESRF Grenoble and Bayreuth Geophysics. The science-industry link is important when looking at exploration tools for oil, diamonds, or mineral wealth. Companies have a vested interest in not making those available yet at the same time they want access to public information. One approach is to propose a time resolution ban. We work with Diamond Trading Company and instrumental companies such as JEOL, Bruker and Philips.

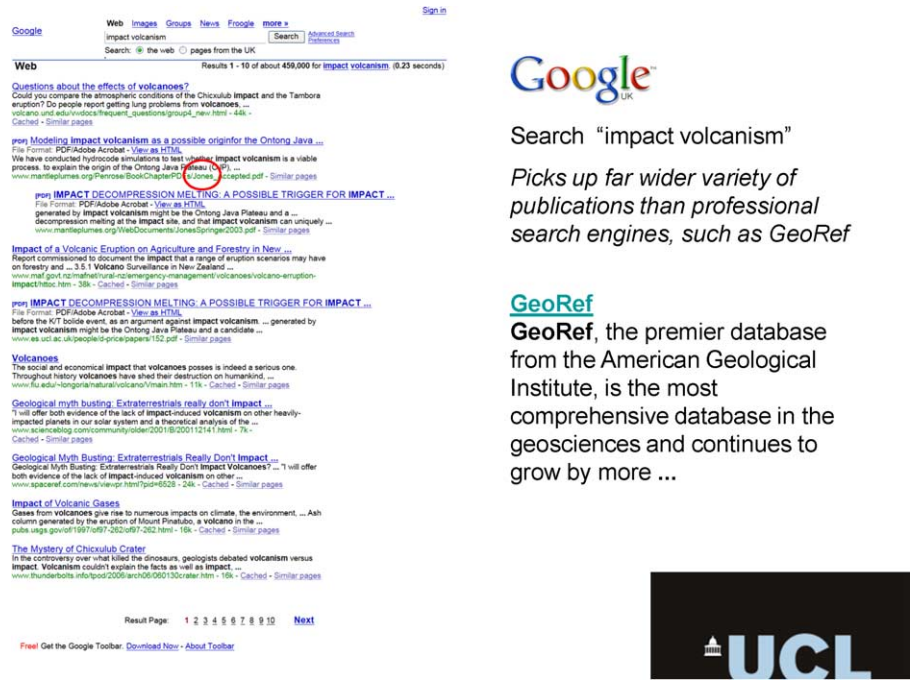


Fig. 4.

Growing sub-disciplines within the earth sciences have supercomputers and access to datasets (such as those for mineral physics, crystallography and climate). Mainstream areas such as mineralogy generate vast amounts of data that are largely not accessible outside of direct ownership. This in turn creates a reliance on journal trawls and older reference books. There are issues surrounding the archiving of mineral data. The tools we use for mineralogy have intermingled with other subjects such as art theft, archaeology, forensics, astrobiology and space. My hope is that more data will be accessible online and available to other people.

What researchers say they want

Michael Mabe, Chief Executive, International Association of Scientific, Technical and Medical Publishers (STM)

At its heart, user behaviour will determine everything that is going to happen and change in the communication environment. To understand how we want to manage information in future we will have to understand how these processes work. The journal is the most important of the scholarly communication metaphors. It was invented by Henry Oldenburg in 1665 in order to satisfy some user behaviour needs of his time. Four key journal functions then were the date stamp, the quality stamp, recording of the final definitive versions of papers and their archiving, and dissemination of results to a targeted scholarly audience.

The objective of the study I undertook at Elsevier was to understand how the motivations and behaviour of researchers has been affected as the Internet reaches early maturity. It was the largest research project of its kind: more than 6,000 respondents (representing all age groups and regions of the world) replied to questions in all subject areas. This was followed up with 70 in-depth phone interviews.

I will share two main aspects of the results in terms of the motivational area for publishing and research highlights that may be relevant to current concerns of the information community and at this conference. The late professor B.R. Coles' study published in 1993 (before the World Wide Web was the dominant force in journal publishing) assessed UK researchers' motivations for publishing. Knowing that people often say what they think their colleagues want to hear, Coles also looked at secondary motivations. Dissemination was found to be the most important of the primary motivations. However, the responses given under secondary motivations reveal some "less noble" aspirations, including teacher funding, career advancement, recognition, establishing of precedence, and dissemination of results; all are consistent with the Oldenburgian functions.

Was this affected by the advent of the Internet? Our 2005 study showed that it was not. The study results in almost every category are similar, not only in terms of the primary/secondary split but also in terms of how secondary motivations are distributed. There are some key differences regarding how motivations have shifted, particularly with respect to recognition and establishing precedence, areas that one could argue have become *more* important in the electronic babble of the World Wide Web. Below are some highlights of the research results.

Funding bodies

Researchers are ambivalent – 68% think that funding bodies have too much power over research conducted. 50% felt some pressure to publish in high impact journals.

Quantity or quality?

With respect to the number of articles published, quality was overwhelmingly felt to be more important than quantity. 70% disagreed with the statement "it is better to publish a large number of papers than a smaller number of quality papers".

A paper I recently published with Mayur Amin (from Elsevier Science) shows that, contrary to popular opinion, productivity rates of unique papers published per year, per unique author are static or declining while the number of unique authors and co-authorship are increasing.

Prestige or niche journals?

From the author's point of view, there is a slight preference for prestige over niche. A significant minority of readers, however, believe that an article's quality is not determined by the journal.

Peer review and refereeing

Peer review is overwhelmingly supported by virtually all the respondents, 88% of whom agreed on the need for refereed journals. (A statistically-significant minority in physics did not agree.) The majority feel peer review improves an article, although there were some sceptics in engineering. 85% were willing to review a reasonable number of their peers' research (from two to 30 papers per year). 40% felt that time constraints prevented thorough refereeing. Other constraints included being sent irrelevant papers, being asked to review poor quality articles, not wanting to review twice for the same journal, and receiving papers from lesser-known journals. (Adrian Mulligan summarised these points in a free white paper available on the Elsevier website.)

Continuous review

42% thought continuous review was important but 32% had real reservations regarding continuity over time of a published source, consistency, and the amount of time required to revise based on continuous comments, some of which may lack relevancy.

The role of the publisher

60% believed the publisher added value. Among the 17% who disagreed there were significant variations: 26% in computer science felt that the publisher did not add value. The same was true for 22% in mathematics. Researchers who had served on funding panels were also more sceptical.

Informal sources

Informal sources (conferences, bulletin boards, emails, etc.) were still regarded as important (21% of the sample disagreed). Informal sources earned greater favour from computer scientists and physicists and less from chemistry and the Life Sciences. On the whole, collaboration seems to have increased, although this is less the case in physics, perhaps because large-scale collaboration has always been the norm.

Reading behaviour

Reading behaviour has clearly undergone change since the pre-Internet studies. A significant minority preferred to do their browsing from home, but e-versions have not yet taken over. Most disagreed with the statement that "an article will only be read if available electronically" (54% in the computer sciences, a large minority in the Life Sciences, physics and astronomy).

Supplementary data

Reflecting some of the concerns we heard earlier, the study revealed strong agreement that all supplementary data should be published. Notably, 75% of respondents want access to others' data while only 52% were willing to share their own – a hesitation based on fears about misinterpretation and/or misuse of data. The most likely primary motivation here is competition.

The article as independent from the journal

26% of respondents always search authors' websites for the full article, although figures were somewhat higher for computer science, mathematics, economics, and for graduate students. About half of those in economics agreed that they would place a full version of the article on their website, whereas only a third in computer science and less than a third in mathematics would do the same. One third of those in computer science would place a final version of the article they are publishing on their website while less that a quarter in mathematics would do the same.

Publication stages

The advent of e-publication has replaced what was once a clear demarcation between published and unpublished with a much less clear series of stages of being published. We now have a non peer-reviewed draft, pre-publication draft, author's manuscript, uncorrected proof, final corrected proof (the "in-press article"), and final published article in paper and electronic, fully-linked format.

Importance and usage issues

The final article is not only the most used but also the most important, although there are significant differences regarding earlier versions and the extent of their use. There is concern that increasing use of the author's final manuscript version after acceptance could create doubt around the saleability of the final article, an issue that was pointed up in a recent study by Chris Beckett and Simon Inger (Scholarly Information Systems) which looked at whether librarians will cancel their subscriptions to the final published version as earlier versions become available. The results indicated that this transition could be close at hand making it a major strategic issue for the information world.

Permanent record

Strong interest in the permanent record, especially with respect to articles published over ten years ago, was demonstrated in economics, social and earth sciences, mathematics, physics and astronomy. In these disciplines older articles are considered classics in their field, provide an overview, avoid the repetition of research, and show that ideas have not changed significantly over time. Some concerns expressed about older versions include the speed with which fields can change and difficulties in accessing ancient archives.

Repositories

Knowledge and awareness of institutional and subject-based repositories was low. Only 5% knew a lot about institutional repositories and 28% a little. 9% knew a lot about subject repositories and 29% knew a little. Repositories have not yet entered researcher consciousness on a global scale. Attitudes to repositories among those who claim to know about them was varied and indicated concern regarding what the purpose of a repository should be, author attribution, repository funding and quality control.

Conclusions

Some behaviours are changing in relation to e-communication utility issues. Researchers are making greater use of technology and are using social networking software. New challenges apply to electronic peer review and global collaboration. That said, most of the fundamental drivers appear to remain unchanged. The four Oldenburgian functions remain extremely present in journals at early Internet maturity. New tools are being used, but they are being used for old purposes. We are unlikely to see these fundamentals change unless the drivers of researcher motivations change radically.

Acknowledgement

The author acknowledges the contributions to this project of Adrian Mulligan at Elsevier and Professor David Nicholas at UCL.

Researchers and discovery services survey and report: an overview

Hugh Look, Senior Consultant, Rightscom

The Researchers and Discovery Services Survey and Report (http://www.rin.ac.uk/researchers-discovery-services) was carried out for RIN, a UK national information and research centre on communications. The aim was to assess use and perception of sections of resource discovery services by

academic researchers in the UK. The results are intended to help determine priorities in future development of services.

About the study

The study was based on a telephone survey of 450 research-related personnel in UK universities, 395 researchers (at PhD level and up) and 55 librarians and information officers across all disciplines. The term "user" was broadly defined. In-depth interviews with postdoctoral researchers supplemented main interviews and were used to assess differences between those who had grown up as researchers in the Internet environment and those who had not. "Resource discovery services", also broadly defined, included bibliographic A&I services, general Internet search services, dedicated guided portals (Intute and H-Net), institutional library catalogues and portals, and libraries and librarians themselves. The draft of the report was reviewed by an expert panel of senior academic librarians, commercial discovery services providers and experts involved in the design of publicly provided resources (JISC).

Key findings

General search engines are the most used; among those Google is used more than any other. Within the library community, the internal library portal was the next most used service. Blogs received limited usage, RSS feeds did not make the list.

Satisfaction with discovery services is high, predominantly among researchers and scientists. In Arts and Humanities there were concerns about gaps in service coverage.

The issue of access (e.g. accessing a document once located) generated greater frustration among researchers and librarians than that of discovery. Another frustration concerned lack of clear delineation between means and ends (between discovery services and what is being discovered).

Most researchers rely on a range of resource discovery tools and select an appropriate tool for a specific inquiry. Researchers in the social sciences appear to use a wider range of resource discovery services than those in other disciplines.

The most heavily used resource discovery sources include general search engines, internal library portals and catalogues, specialist search engines and subject-specific gateways.

The pattern of researchers' named discovery resources is expressed by a long tail; a very few resources – Google, Web of Science/Web of Knowledge and ScienceDirect – are named by a large number of researchers. Researchers also showed reliance on interfaces such as Athens (an authentication system) which they inaccurately identified under services.

Among the range of resources found through use of discovery services journal articles are the most important. Virtually all researchers (99.5%) rely on the journal article as a key resource. Over 90% also use chapters in multiple author books, organisation websites and individual expertise. The next most cited resource – monographs – is mentioned by only 32% of researchers.

Peers and networks of colleagues are shown to be extremely important for virtually every type of inquiry. Research colleagues feature as important providers of information about resources and tools and new services, and this is particularly the case for postdoctoral researchers. Some researchers use email listservs, however online social networking services have been less popular. Colleagues are relied upon for locating individuals, initiating research, discussing research funding and locating data sets.

The majority of researchers work by refining down from large sets of results. Surprisingly, researchers were more concerned about missing important data than they were about the amount of time spent

locating information. Concerns were also expressed about being overwhelmed by email, and in every discipline researchers bemoaned the number of irrelevant results delivered by general search engines.

With respect to emerging tools, blogs were shown to be little used. A majority of researchers (62%) obtain regular information updates and alerts from services pushing information to their desktops, and email is the preferred tool for this (not RSS feeds). A smaller number use alerts on funding sources from research councils or specialist services and some specifically cite Zetoc (an electronic Table of Contents service from the British Library) as a useful service. Sources for keeping up to date include journals themselves, email alerts, conferences and conference proceedings, among a wide range of "other" sources which will not be discussed in detail.

The focus of library activity has shifted. Library support is now being delivered more often through services provided than personal contact. Librarians' and researchers' views diverged on a number of key issues including quality of discovery services, availability of resources, and gaps and problems. Researchers do their own searches in the vast majority of cases. Librarians overrated the importance of datasets to researchers. Librarians used general search engines far less frequently than researchers. Librarians perceived researchers as conservative in their use of tools and were concerned that they were not reaching all researchers with formal training. Researchers did not perceive this to be a problem.

Specific gaps in provision included access to foreign language materials, lack of distinction between actual sources and discovery services, difficulties in locating specific chapters in multiple-authored due to lack of general indexes, and too-short backfiles of journals.

A plea for "one stop shops" was made across the board. Researchers have come to agree that "the more digital, the better". Most expressed concern about not having access to a sufficient number of digital resources. Problems cited included institutions not subscribing to the full text of the e-journal, lack of journal and source text digitisation, and overly short electronic backfiles.

The data showed fewer differences between experience cohorts than one might expect. Regarding frequent and regular use by experience, age did not play a significant role, although the younger group stands out clearly in the use of blogs. Differences between disciplines are somewhat more marked. Researchers in the Life Sciences make more use of their colleagues than in other disciplines. In the Physical and Life Sciences, researchers tend to use general search engines more than average. The library portal is used more frequently by Arts and Humanities lecturers.

Google is not being used for mission-critical applications. Rather it is relied on, often in combination with other tools, to locate organisations and individuals, references, or to research a new area.

A wide range of resources are used including bibliographic databases, Google, internal portals, Web of Science and Web of Knowledge. The category "other" (46%) reveals a wide variety of discipline-specific resources.

It can be noted that the boundaries between resources themselves and discovery services are increasingly permeable, a trend that is likely to continue as new forms of content aggregation are developed.

Rethinking user behaviour in a Web 2.0 world

David Hoole, Marketing, Nature Publishing Group (NPG)

My aim is to rethink user behaviour in a Web 2.0 environment and place it in a development context. 2006 was a significant year in the information industry with the emergence of Web 2.0. The tools to host and navigate content on the Web are largely in commercial hands. As scholarly publishers we have dealt with user-generated content for many years. Developments are perhaps not as radical as they seem, and the current environment may be one in which we can prosper.

User behaviour is changing. I would emphasise the importance of users as original sources of content and publication ideas, the tendency of journals to develop in particular directions that may have not been foreseen by publishers, and the community-building aspect of traditional academic journals. During the early 1990s, tracking print journals was increasingly difficult. Search culture took hold in the late 1990s with Yahoo and Google. A key development occurred with the movement of A&I services onto the Web. Among these, PubMed is the most widely used. Increases in submission rates demonstrate how the Web connected and expanded scientific communities and has benefited publishers. In recent years, the focus has been on search and lowering the barriers to Web publishing.

I will focus on search behaviour, an area that is often oversimplified. It is important to highlight different kinds of search activity. There is no such thing as a "one stop shop". Users always want more and rely on multiple information seeking episodes. Disappointingly, most usage on publisher websites is the retrieval of PDF files. While we may see thousands of searches, we see many millions of article downloads.

Not much is new in Web 2.0. We should be concerned about the ever-increasing power of Google, especially as our own brands appear weaker by comparison. A key difference is that many more people are involved, the result of broadband access. There is nothing particularly unusual about users' behaviour – just the extent of it. As information professionals, we are well placed to bring some order to the supposed chaos. We need to look at user behaviour in new ways to identify new opportunities.

Although few of us do this, search terms could be used as indicators for product or content development. Search should not be the sole guide as it could lead to overlooking important information. At the same time, it may be a self-fulfilling prophecy – as more is published in a particular area (e.g. stem cells), more users will come to a site looking for stem cell content which will result in that search term being used all the more.

Could we use key words to help map Web resources? For example, if a particular word is used (and found) in a number of papers, is that an indication of the resource's coverage of the topics? I compared topic markers to genetic markers. Then, using search terms, I mapped a number of resources. The terms I used are grouped into physical and life sciences, biomedicine, and social, political, and technological terms, and covered nine resources including Science.gov, Connotea, Wikipedia, nature.com, Blackwell's Synergy, PubMed Central (PMC), ScienceDirect, Scopus and Google.co.uk.

The key words were inspired by search term lists, hot topics and some personal choices. It is not clear how different selections would have impacted the results, but the important thing is to use the same set of terms across the databases to be assessed. Pitfalls in search include problems with certain phrasings (RNAi) and formatting (using underscore as a space). To assess the impact of this I did searches on RNA and RNA inhibition. Search results cannot be used to map a resource where the true word count is not available (this was particularly the case with Science.gov which, because it delivers a set number of results, ultimately became worthless as a measure and was excluded from further analysis). Results tend to change daily, and Google, which is operating in a league of its own in terms of high search results, demonstrated some shocking variations. It should also be noted that certain words like "brain", "AIDS", "cancer" and "policy" have a high frequency in all literature. We can simplify by just looking at the life sciences and excluding Google and one or two other websites.

We found a strong pattern showing that the major Life Sciences resources have a similar balance of content (Fig. 5). Scopus has the most items so one would expect it to be the largest resource. PMC and ScienceDirect have fewer items, but ScienceDirect remains the larger resource. Nature.com and Blackwell's Synergy are much smaller resources in terms of reference material although this is not an indication of quality or overall value of the resource.

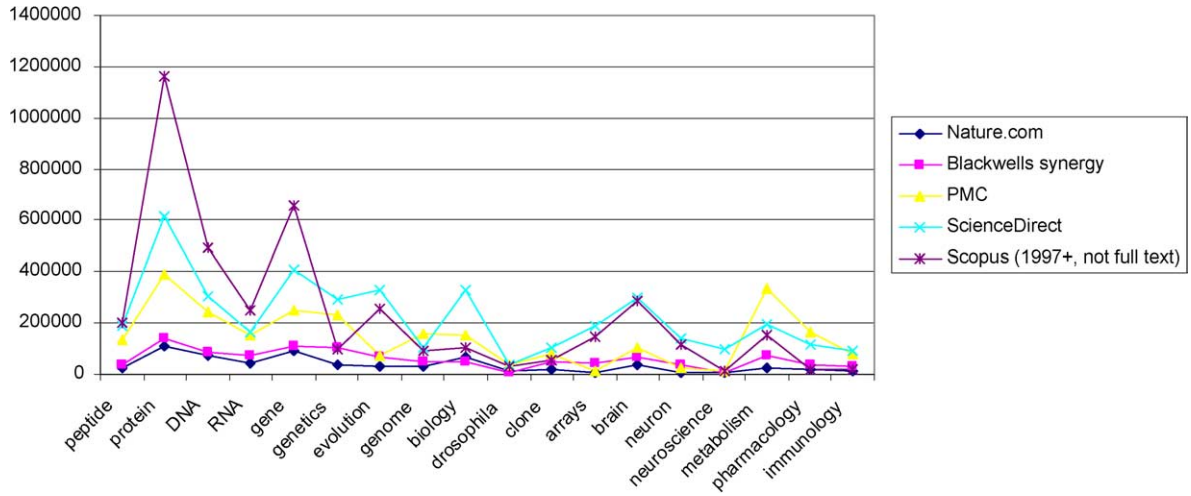


Fig. 5. Life sciences.

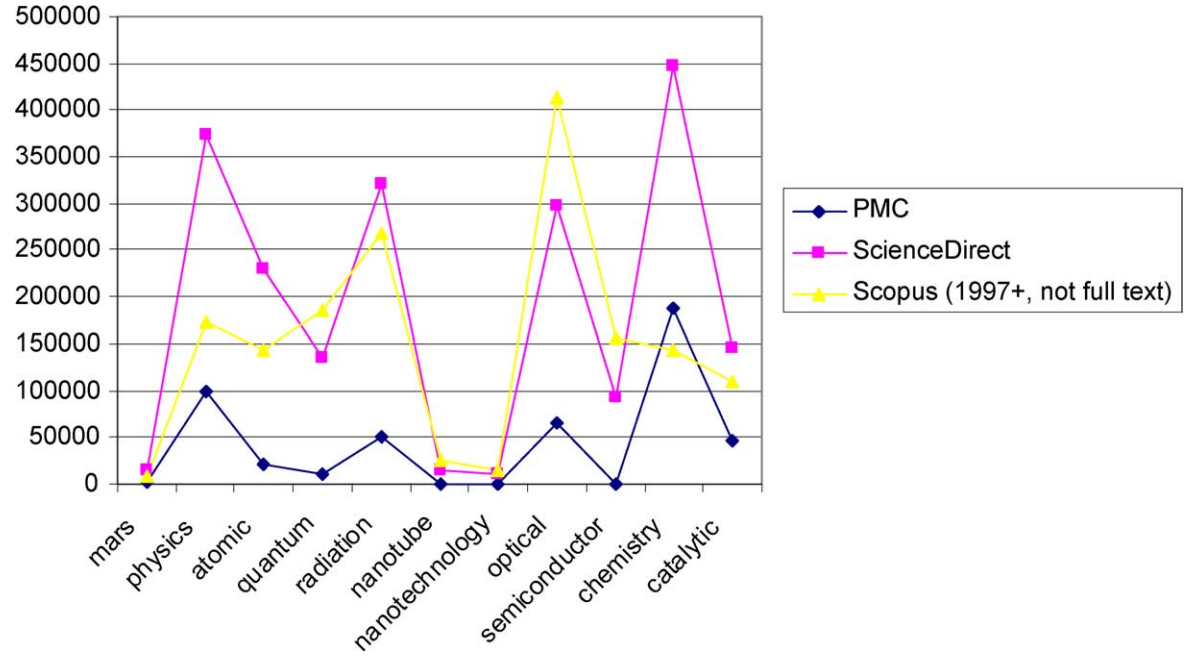


Fig. 6. Physical sciences.

In physical sciences (excluding Google, see Fig. 6) there are expected results and ScienceDirect has more content than PMC. ScienceDirect and Scopus are similar. The reasons for the odd low Scopus chemistry point are unclear, although it could reflect Elsevier's strength in chemistry.

Looking at biomedicine (Fig. 7), I have included a logarithmic graph showing that Connotea remains a relatively small resource. One notes the close correlation between Scopus, ScienceDirect, PMC, Blackwell Synergy and nature.com. The standout feature is Connotea's strength in bird flu, something we have known for some time (by Google page ranking) but which we have only now been able to show in

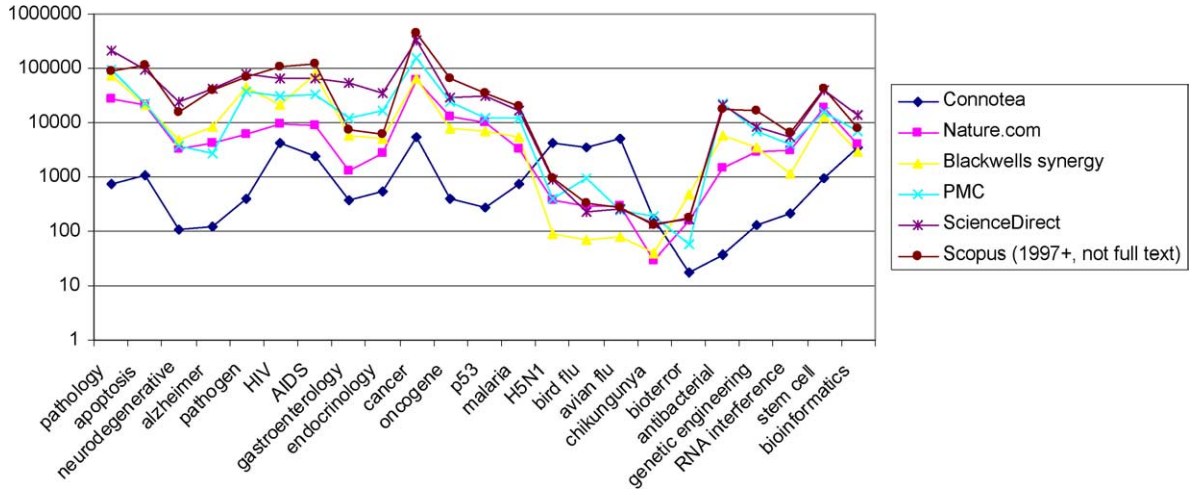


Fig. 7. Biomedicine.

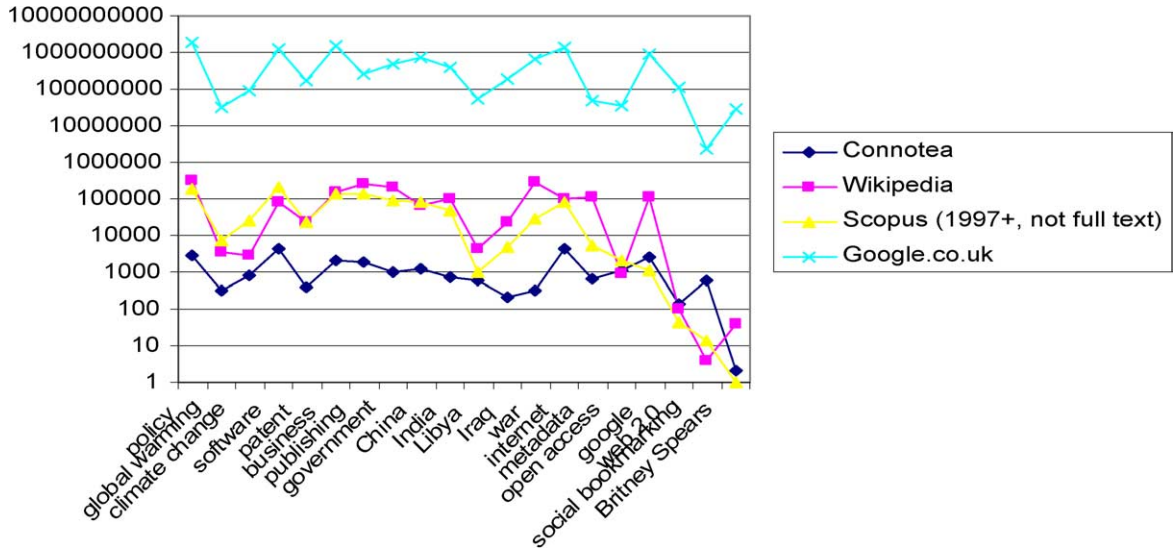


Fig. 8. Social/political/technological.

a graphically quantifiable way.

Chikungunya is a disease characterised by high fever and paralysis from which up to 25% of children in post-Tsunami orphanages were suffering. Interestingly, Connotea has almost as many reference to Chikungunya as PMC and ScienceDirect, despite being a much smaller resource – a possible reflection of Connotea's strength in avian flu. The site has expanded its coverage of other disease topics as overlapping communities post and tag related articles. It may also be a reflection of the oft-quoted suggestion that user-generated content is topical, ahead of the curve, and several steps ahead of traditional publishers. In a rapidly changing world, early highlighting of important new topics is exceptionally valuable and a potentially huge aid to publishers and scientists.

Looking briefly at social/political/technological topics (Fig. 8), we find that Wikipedia and ScienceDi-

rect are well correlated. Scopus does not reference Google as much as Wikipedia, Google and Connotea (and why would it?). Connotea, a social bookmarking tool, has many references to social bookmarking (issue of self-referencing). It might be easy to conclude that Connotea is just a random collection of topics, linked only by their topicality or users' self interests. But traditional journals have done the same and have been guilty of self-citing and specialising in only parts of their declared scope.

We may need new tools to evaluate and analyse Web 2.0 publishing, but we have always worked with user-generated content and should be well positioned to add value. Many scientific challenges lay ahead. If we can monitor hot topics and respond quickly as publishers, we may be able to aid in a variety of transitions. There is an opportunity for encouraging better indexing and to move toward user-generated thesauri to help eliminate the problems and pitfalls regarding search terms and user-generated tagging.

SESSION II
METRICS

The great impact factor debate

Jim Pringle, Senior Vice-President, Thomson Scientific

I will approach the issue of impact metrics using insights from historical studies and geology. Rather than making another critique of the impact factor, I want to look with a fresh eye on the debate itself and to what it says about the underlying trends affecting the research community, behaviour of researchers and their many roles, and that of the organisations that serve, judge, and employ them. Today, citation-based impact measures are fiercely debated – I will not focus on this here but to say that it is a debate that reflects deeper shifts in goals of research; impact and other metrics will evolve and adapt to the changing nature of the research community.

Future historians, will look back upon 2005 as the year the journalists discovered the impact factor (judging from the number of documents containing "impact factor" and "journal" indexed in Web of Science in January 2007 (see Fig. 9). Attendance at bibliometric conferences is breaking records. Fields such as bibliometrics and scientometrics are undergoing a renaissance. I would suggest that the journal impact factor, a metric of particular interest to the research community, is under intense pressure and has become a general proxy for all research value. And the debate about journal impact factor has become a proxy for a debate about all quantitative measures of research value, the fundamentals and causes of which are to be found beneath the surface.

Much of what we have discussed so far has focussed on broader trends such as research funding, university productivity, redefinition of researcher community and communications, globalisation of publishing, library purchasing patterns, and availability of metrics. Increasing competition for funding among the research community creates pressure on the interaction between evaluators of grant programmes and researchers applying for those grants. There is increasing emphasis on outcomes with respect to university productivity. The Internet presents new ways of collaborating and competing as reflected in many of the redefinitions of researcher community and communications that we have heard today. More journals – all of them seeking recognition – are being published in new areas of the world. Within the library community, more attention has been paid to evaluating content in new ways. Also of interest is the sheer availability of metrics to be debated. If open access is a question for scholarly publishing of research articles, it also affects global use of the impact factor.

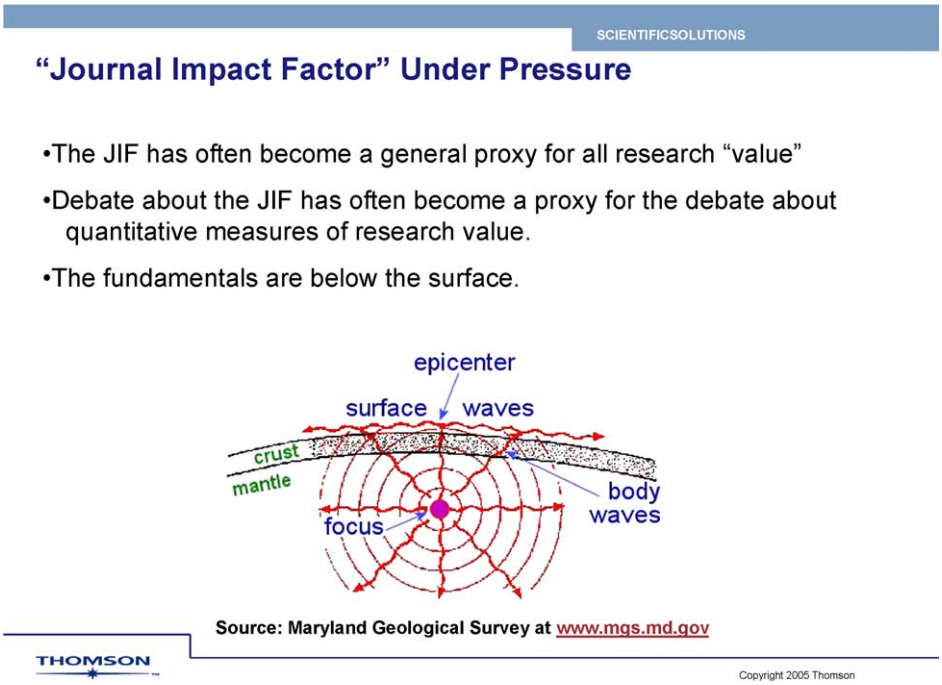


Fig. 9.

These underlying trends create fault lines that affect the question of value in four key areas: publishers demonstrating the value of journals and processes (validation and prestige); funders assessing value for decision making (how has our money been spent?); the open access community demonstrating the value of new models of publishing (are open access journals as good as other journals?); and researchers demonstrating the value of individual work. We are now looking at open peer review and social networking as new models for evaluation and judgments about value.

The debate is driven by the question of value. I will not discuss or defend which impact metrics are the right ones to use but would speculate that new conceptions of the research community are merging along the fault lines alluded to earlier. For example, at the intersection of the open access and the research communities arise questions of participation, peer review, and potentially new models of evaluation. At the intersection of funding and open access communities lies interest in institutional repositories. Within the communication between publishers and researchers there is a community of interests. The publishing community is increasingly trying to create new ways of linking the journal publisher with the author.

The debate on metrics gravitates toward value and new ways of assessing it. Thomson Scientific has always listened closely to the publishing community about how the impact factor is calculated. In certain areas of research communication, there has been a call for new types of indicators and metrics. A wealth of new experiments, such as new ways of calculating the impact factor (appearing at a rate of about one every two months) are emerging from within the publishing and funding communities as well as the bibliometric community that serves them. New indicators are emerging from support research and communication, on behalf of publishers, and from other communities. These focus on article downloads, article citations, local publication and research activity and Web linkages.

We are currently in era of experimentation and will likely move from impact factor as a proxy for a complex debate about value toward a more integrated decision support environment. Within Thomson

Scientific, we are focusing more on what are the right tools for the right job and replacing a "one-size-fits-all" approach. In 2006 we launched *Journal Use Reports* which bring usage, publication activity and citation metrics to the librarian and the provost. We are working with ScholarOne and Manuscript Central to place citation data in a context that will allow editors and reviewers to benefit from the right information at the right time. We are incorporating individual indicators and analysis into the Web of Science to respond to issues at the upper levels of the funding administration. It is likely that as a statistic, the journal impact factor, perhaps supplemented by other statistics, will remain on the scene as long as peer-reviewed journals are published.

Moving from use to users, and then on to outcomes

Professor David Nicholas, Centre for Information Behaviour and the Evaluation of Research (CIBER), UCL Centre for Publishing

I will discuss what scholars actually do (not just what they want). CIBER has been researching user behaviour using deep log analysis for about five years and built an evidence base of some 6 million users and millions of usages. This may represent the first time anyone has longitudinally linked up a database over a period of years. The programme is called Virtual Scholar.

We are constantly rolling out digital initiatives to the scholar/researcher. Monitoring and evaluation is needed to feed back to and modify the existing system. We are in the age of the (disintermediated) digital information consumer – promiscuous, volatile, independent and anonymous. Most of our searching is not done from the library. Is there a danger of libraries decoupling from the information community?

We still do not adequately understand the user and user metrics. We need evidence of usage, to move from hits and downloads to users through *deep* log analysis. Only when we have *user* data can we move on to impacts and outcomes. Deep log analysis involves mining logs to the *n*th degree and relating digital fingerprints to the users who made them with respect to age, gender, and other characteristics. In this study we have even related usage to attitudes towards scholarly publishing issues such as peer review. But is there a way to measure to what extent increasingly expensive access to information results in improvement?

The good news is that the appetite for scholarly information is growing. The bad news is that much of this activity constitutes "ferreting around". People searching from home are perceived by libraries as not using the library's services. A high proportion of users view only a few pages from a vast number available and a high proportion do not return. This creates huge volatility in the system, but what does it mean? People do not read online. They spend more time viewing shorter articles than longer ones. The longer the article the greater the chance that it will only be viewed in abstract. Many people say they are going to download an article and read it later, but do they? A picture of a consumer audience is emerging.

It is worrying that 'number of downloads' is a key metric exchanged by librarians and publishers and used as a measure of success. No one is considering whether this is a flawed metric because everyone is using the currency and wants it to remain in circulation.

People who come in via search engines do different things: look at older articles and view more subject areas, journal titles and articles in the abstract. We have found enormous differences among users according to subject field, academic status, geographical location and gender. No one asks about gender, even though in marketing gender is a major variable. How many studies on scholarly communication actually factor this in? Among the many interesting findings was that women return more often than men

to the same site. Staffers spend more time searching while students spend more time viewing articles. Older people are more likely to have one-view sessions. Mathematicians appear to be far more regular visitors than engineers. We can combine this information to create a profile on what a particular community does. The data in itself is not significant, however the issues that it raises are. Once we can relate user activity to user characteristics and identify differences between them, we can start asking why and seeking explanations. This bring us to impacts, performance and outcomes.

There is an unwritten assumption that (digital) information consumption does you good otherwise it would not be so expensive. But this needs to be demonstrated. We are on our way to doing this. Although we have been talking about users for years, very little progress has been made. Can we take it further? Can we move beyond a technological fix? The main challenge is to accommodate the concept of the digital information consumer. We need to address questions that arise from the logs and see what they tell us about outcomes.

What I have been discussing was not derived from the opinions/perceptions of small and unrepresentative samples. It is something that has been building up over five years. We looked at Blackwell Synergy, ScienceDirect, OhioLink, OUP Journals, and others and a basic pattern has emerged. We are seeing a form of consumerism. These are big questions that affect our world in a substantial way.

Making statistics useful

Dr Peter T. Shepherd, Director, Counting Online Usage of Networked Electronic Resources (COUNTER)

One of the most significant changes in publishing has been the development of tools and metrics to understand user behaviour. I will discuss Project COUNTER, its background, current status, and how COUNTER statistics can support the development of other metrics. In conclusion, I will ask to what extent our products are successful and valued, and what they tell us about user behaviour.

The COUNTER project was an industry initiative begun in 2002 which saw the development of two codes of practice to improve quality, comparability and credibility of vendor-based user statistics. Main elements include:

- Agreement on terms used among publishers. COUNTER established a clear set of definitions among publishers
- Establishment of specifications of the usage reports which vendors deliver to customers (appearance, timing of delivery)
- Set data processing guidelines (e.g. date/time filters to address multiple requests)
- Implementation of an independent audit of reports and processes, creation of a compliance procedure
- Representation of the governance community with respect to COUNTER's governance.

Current status

Two codes of practice have been produced. The first, for online journals and databases, was initially released in January 2003 and an improved and validated version became available in January 2006. It has been widely adopted with 60% of Science Citation Index (SCI) articles covered by the COUNTER code of practice. In 2006, it was extended to cover online books and reference works. This proved more challenging to develop as it was harder to define the most meaningful content unit. The compliance rate is currently low.

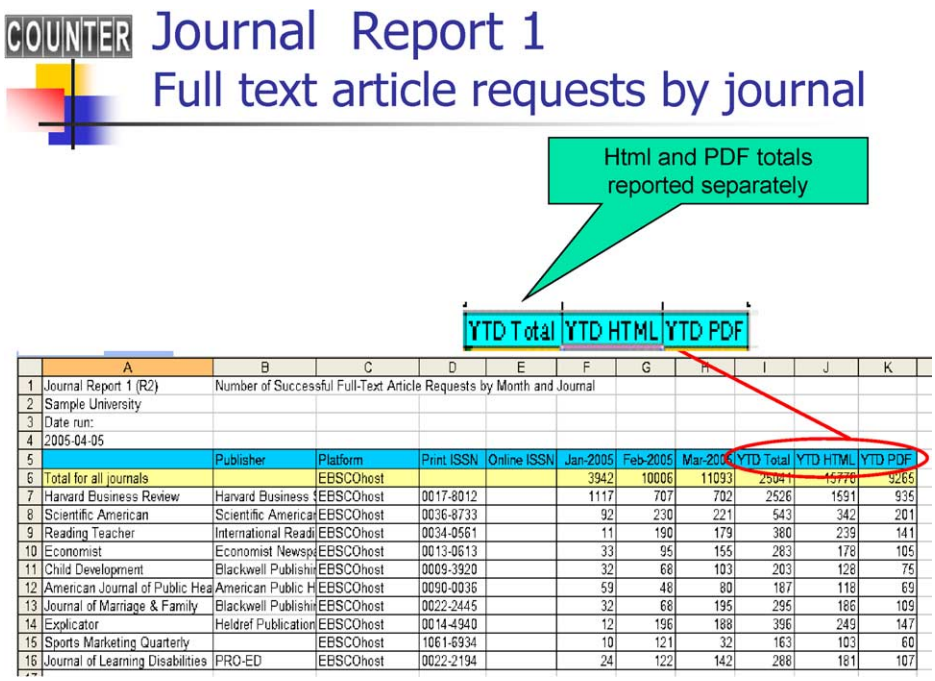
Local and global metrics

I will look at one report as an example. General Report 1 ("Full text article requests by journal" – Fig. 10) has led to the creation of two types of metrics: local and global. Local metrics are useful to individual libraries and consortia and can be consolidated at the individual journal, collection or publisher level. We are also looking at more global measures that could be of use to authors, funding agencies, libraries and publishers.

In 2005, JISC conducted a national overview of online journal usage which led to the creation of a metric (the basis of which was the number of full text article downloads) applicable to all academic institutions in the UK. COUNTER looked at data for 17 participating institutions that were analysed in terms of usage range of journals (high, medium, low), price bands and subject categories. Among the metrics derived, the most important was cost of full text article requests and cost per user. Among other conclusions were that the cost of downloads varied greatly across institutions and publishers' Big Deals have proven a good deal. (A summary of the report is available on the JISC website.)

The impact factor (as discussed by Jim Pringle) is a well-established measure endorsed by funding agencies and researchers. It does not cover all fields of scholarship and tends to reflect the value of journals and research communities. It is subject to overuse and in some cases may disguise the real value in citation data. A usage factor should cover all online journals, reflect the value of journals beyond the research community, and have a clear relationship with the data from which it is derived.

United Kingdom Serials Group (UKSG) has sponsored a project to assess the feasibility of developing and implementing something that might be called the journal usage factor. The purpose of the two-part, survey-based project will be to determine support from authors, librarians, and the publisher community for such a measure and to ascertain practicality of implementation. Some questions to be addressed:



	A	B	C	D	E	F	G	H	I	J	K
1	Journal Report 1 (R2)	Number of Successful Full-Text Article Requests by Month and Journal									
2	Sample University										
3	Date run:										
4	2005-04-05										
5		Publisher	Platform	Print ISSN	Online ISSN	Jan-2005	Feb-2005	Mar-2005	YTD Total	YTD HTML	YTD PDF
6	Total for all journals		EBSCOhost			3942	10006	11093	25041	15776	9265
7	Harvard Business Review	Harvard Business S	EBSCOhost	0017-8012		1117	707	702	2526	1591	935
8	Scientific American	Scientific American	EBSCOhost	0036-8733		92	230	221	543	342	201
9	Reading Teacher	International Readi	EBSCOhost	0034-0561		11	190	179	380	239	141
10	Economist	Economist Newspa	EBSCOhost	0013-0613		33	95	155	283	178	105
11	Child Development	Blackwell Publishir	EBSCOhost	0009-3920		32	68	103	203	128	75
12	American Journal of Public Hea	American Public H	EBSCOhost	0090-0036		59	48	80	187	118	69
13	Journal of Marriage & Family	Blackwell Publishir	EBSCOhost	0022-2445		32	68	195	295	186	109
14	Explicator	Heldref Publication	EBSCOhost	0014-4940		12	196	188	396	249	147
15	Sports Marketing Quarterly		EBSCOhost	1061-6934		10	121	32	163	103	60
16	Journal of Learning Disabilities	PRO-ED	EBSCOhost	0022-2194		24	122	142	288	181	107

Fig. 10.

should the usage factor be based on the number of full text downloads? How to count article numbers? The creation of a global figure requires a process to consolidate, calculate, audit and report on usage factors. There is also a question of how to define total usage. We hope to complete the study and deliver a report at the UKSG conference at Warwick University in April 2007.

It is early for conclusions but I will share some preliminary feedback and questions:

- The COUNTER user statistics are not yet sufficiently robust but will be in the future
- There is frustration about the lack of comparable data on journals in the public domain
- Should the items covered in the study be restricted to articles?
- Many journals have significant usage in print; how is this to be addressed if only covering online usage?
- There is a diversity of views on what should be the specified usage and publication period covered
- Like impact factor, the usage factor can be manipulated; caution is advised
- As journals are increasingly bought as large packages, will the journal be a meaningful concept in the future?
- Access is increasingly via Google and other search engines.

Two measures with different limitations are probably better than one. COUNTER is a technology-based project and will require modification and/or updating. Some issues being addressed at the moment include:

- How do interface and user behaviour affect these statistics? Publishers, responding to librarians' need for purchase reports, are providing a separately purchasable digital archive. We are now developing COUNTER reports specifically aimed at archives
- Institutional repositories are increasingly using journal articles, however we do not currently have any institutional repositories that are compliant; how do we sell this to them?
- What do you count? How to you address hybrid journals? Partial open access journals?

Customer satisfaction and performance metrics

Dr Elliot R. Siegel, National Library of Medicine, National Institutes of Health, USA (acknowledgment to Fred Wood)

In this presentation I will provide an overview of a multidimensional approach to Web evaluation, and offer an in-depth look at online user surveys, with particular reference to the American Customer Satisfaction Index (ACSI). I will conclude with an evaluation of this performance metric, as carried out over an 18-month period with 60 websites that were deployed at the National Institutes of Health to meet a range of information needs by biomedical researchers, clinicians, and the general public.

Access to most databases is typically via online websites. We have an imperative to understand who these users are, what they are getting from us, how satisfied they are with our product, and how they are using the information obtained. The methods used to undertake this should vary based upon the stage of the website lifecycle (in development, operational, or under improvement) and we should be able to triangulate and integrate evaluation data from different sources.

Figure 11 shows a multidimensional approach to Web evaluation. Usability testing is an important means of obtaining feedback from prospective users. Usage data include superficial data, weblog and deep weblog data. User feedback encompasses online surveys, syndicated surveys, focus groups – the means to speak and interact with your users to get feedback on functioning in real time or a later phase.

Table 1. Criteria for selecting Web evaluation methods.

	Web site life cycle stage		
Evaluation method	**Development**	**Operations**	**Improvement**
Usability testing			
Heuristic or expert review	✓✓		✓✓
Usability lab testing	✓✓✓		✓✓
Informal usability feedback	✓		✓
User feedback			
Online internal user survey		✓✓✓	✓✓✓
Online external user survey		✓	✓
Focus group	✓✓	✓✓	✓✓✓
Nationwide syndicated survey	✓	✓✓	✓✓
Unsolicited user feedback		✓✓	✓✓
Usage data			
Web log data analysis		✓✓✓	✓✓✓
Internet audience measurement	✓	✓✓	✓✓
Web and Internet performance	✓	✓✓	✓✓

✓✓✓ = Very important; ✓✓ = Moderately important; ✓ = Less important;
No check mark = Generally not applicable or not important.

Fig. 11. Web life cycle concept. (From Wood, Siegel, et al., A practical approach to E-government Web evaluation, *Information Technology Professional*, May/June 2003.)

Web and Internet performance data look at evaluation from the perspective of the server that provides access to a website and from the surfer downloading. Different methods work better for different needs.

Online survey metrics have improved greatly since the late 1990s when they were costly and lacking in standard methods and benchmarks. ACSI, which we started using in 2003–2004, provides greater value added by allowing for a continuous collection of data in real time, rigourous standardised survey methodology, randomised intercept and a rolling sample that changes on a monthly basis, a series of standardised questions with the opportunity to create optional custom questions, and extensive benchmarking of results. Further, ACSI allows you to:

- Measure the satisfaction of your website visitors on a continuous basis. Are you meeting visitors' needs? Exceeding their expectations?
- Measure the impact of changes (planned and unplanned)
- Identify the potential impact of changes made to your website that could lead to improvement in overall satisfaction scores
- Benchmark performance of your site with similar sites.

Using this methodology is not necessarily simple and it is important to have people at your end who are capable of understanding the continuous stream of data produced. It is possible to review the results in real time online, and every four to six weeks you receive a written report and can review findings with a team that meets with your website team. A quarterly briefing is given to ACSI customers in the Washington area.

Figure 12 ("Illustrative Data Reporting of Survey Results") is the most important slide of this talk. Element scores (for content, functionality, image, etc.) are shown on the left; each element receives a satisfaction score in its own right. The makers of the survey explain that the underlying statistical methodology enables you to predict an increase in satisfaction linked to changes in each individual element score. It is therefore possible to identify those specific areas where modifications will lead to greatest overall satisfaction with the website (and to spend your money where it will *really* count).

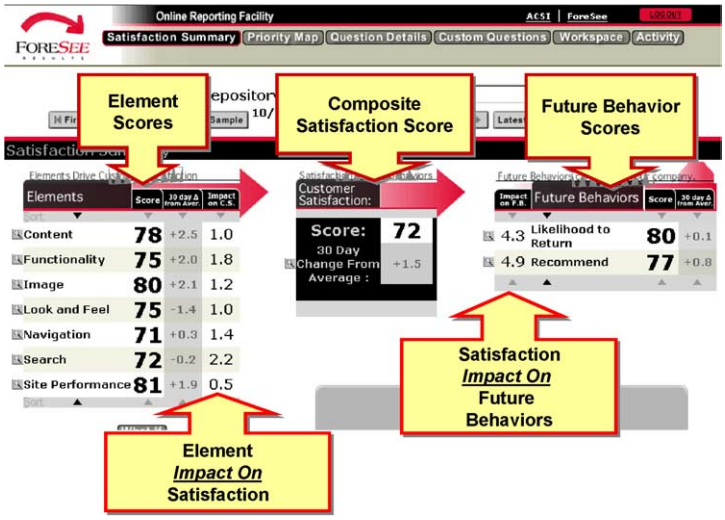


Fig. 12. Illustrative data reporting of survey results. (Source: ForeSeeResults, Inc.)

The ability to create custom questions enables you to focus on specific queries regarding frequency of visits, type of visitor, and primary means of finding the site, and subsequently to triangulate and co-ordinate those data with those from your weblog analysis. If these do not match up, there is a problem.

The evaluation started with 60 websites but not all of them had sufficient traffic to generate the data needed. Of these, 55 sites were active well into 2006 and 42 sites collected enough survey data to generate ACSI scores. We conducted an evaluation of the evaluation to assess the use and value of ACSI to website teams. Our findings showed:

- A majority of respondents strongly or somewhat agreed that the ACSI scores and custom question results were useful
- A majority cited one or more key uses of the ACSI data and their intention of using ACSI data in the next redesign
- About three-quarters cited one or more types of site improvements planned using the ACSI data
- About two-thirds strongly or somewhat agreed that they were satisfied overall with the ACSI.

Another major goal was to evaluate the importance of the ACSI to NIH as a whole. Here it was found that the project:

- Greatly increased the focus on measurement of customer satisfaction with NIH websites
- Encouraged a user-cantered approach to NIH website design and improvement
- Strengthened the network of NIH website professionals
- Provided opportunities to share experiences, lessons learned, and informal mentoring.

Conclusions

Where the ACSI was useful and successfully deployed it tended to be correlated with timing of the survey. For example, ACSI was found to be more useful in instances where a site was planning on a redesign. Managerial support and sufficient financial and human resources to handle the data were all important. Difficulties associated with the deployment of the survey were related to low traffic websites,

insufficient volume for valid online results, on intranet sites, and where staff and management were sceptical.

The ACSI, like all online surveys in the Web environment, has a relatively low response rate (typically in the range of 4–8%). ACSI uses random intercepts and several crosschecks to minimise non-response bias, but this remains an issue that warrants greater research attention to understand whether surveys in the Web environment really give us the answers we seek. Based on the NIH experience, the ACSI would seem a particularly applicable medium for assessing high traffic websites in any country and in fields of science, technology and medicine that have a significant "public" user base.

SESSION III
FUNDING

Providing support for UK biomedical authors and users

Richard Boulderstone, Director eStrategy, The British Library

I am here to try and figure out what writers should do. I will discuss the British Library's strategy and approach with respect to STM and look at the current biomedical and health research environment. I will also share some early results of CIBER and talk about a project we have just started called UKPMC.

The British Library is the national library of the UK and has some 250 years of continuous collection. Our mission is defined as "helping people advance knowledge to enrich lives". We run the world's largest document supply business, although volume has declined over the past decade primarily as a result of the Big Deal. Nevertheless, we still distribute about 2 million articles a year on that service (about 80–90% of the material from STM journals). About a year ago we looked at revising the overall strategy for the library – or "re-defining the library" – in the face of a range of challenges, many of which have been discussed today.

Regarding overall strategy, we think of things in disciplinary chunks: Arts and Humanities (we think we understand them – they come to the library and look at our "stuff"); social sciences (confusing to us because they want a combination of textual sources, databases, and data sets and applications); and STM (unclear what is going on as STM reading rooms are not much used). Researchers are voting with their feet and coming here less. This is a huge challenge for us.

We discussed whether we wanted to be primarily an Arts and Humanities library or to try and represent the whole sphere of disciplines across the research community. We decided to maintain the coverage and to span all disciplines, but this represents a major challenge in the STM area. We conducted research that showed many people in the STM disciplines did not know who we were, particularly as librarians mediate the supply service. We found our relationship to researchers to be relatively limited and saw a need to become more user focused.

For our mostly text-based collections, the core strengths of which are size, scope and scale we need to focus on researchers' needs and behaviours and what products and services we can provide on top of this collection.

We focused on biomedicine and health as a means of segmenting the STM area (too large to take on as a whole). Figure 13 ("Correlating Discipline Areas with Researcher Numbers and Growth") shows growth experienced by undergraduates and graduates in the UK. The biomedical area, which shows large numbers of researchers, is one in which we can have an impact. Research councils and their funding priorities (Fig. 14) also have a direct impact on biomedical areas.

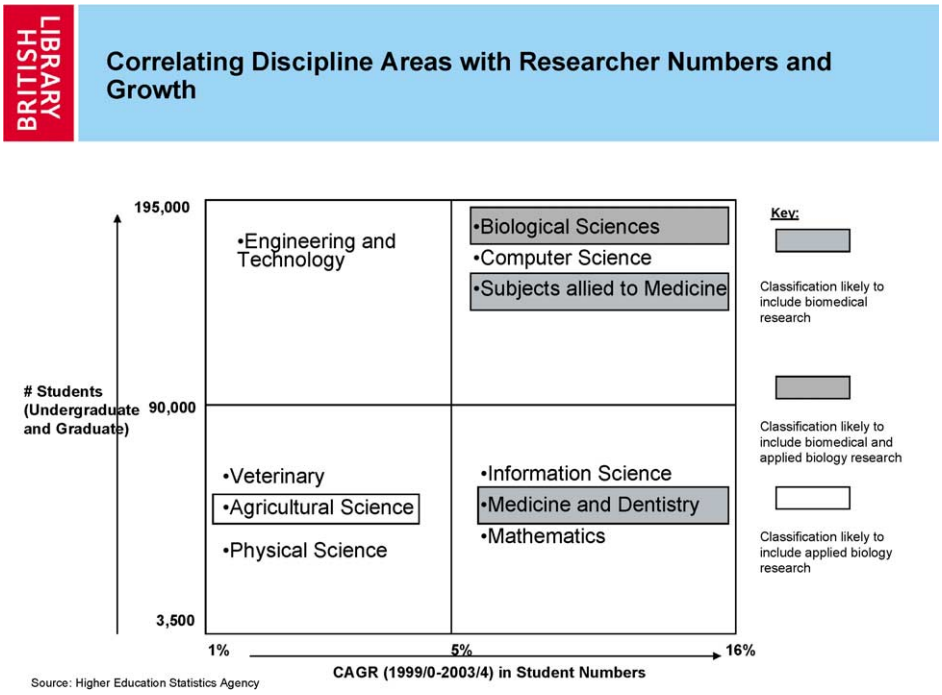

BRITISH LIBRARY
Correlating Discipline Areas with Researcher Numbers and Growth
195,000
90,000
3,500
Students (Undergraduate and Graduate)
•Engineering and Technology
•Biological Sciences
•Computer Science
•Subjects allied to Medicine
•Veterinary
•Agricultural Science
•Physical Science
•Information Science
•Medicine and Dentistry
•Mathematics
Key:
Classification likely to include biomedical research
Classification likely to include biomedical and applied biology research
Classification likely to include applied biology research
1%
5%
16%
CAGR (1999/0-2003/4) in Student Numbers
Source: Higher Education Statistics Agency

Fig. 13.

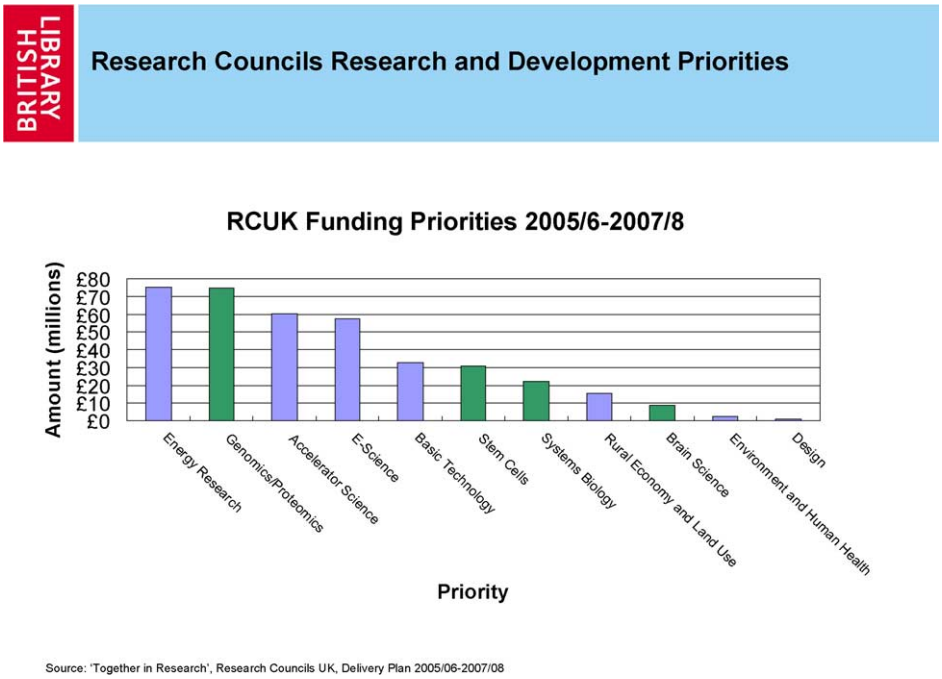

BRITISH LIBRARY
Research Councils Research and Development Priorities
RCUK Funding Priorities 2005/6-2007/8
Amount (millions)
£80
£70
£60
£50
£40
£30
£20
£10
£0
Energy Research
Genomics/Proteomics
Accelerator Science
E-Science
Basic Technology
Stem Cells
Systems Biology
Rural Economy and Land Use
Brain Science
Environment and Human Health
Design
Priority
Source: 'Together in Research', Research Councils UK, Delivery Plan 2005/06-2007/08

Fig. 14.

What are biomedical/bioscience researchers experiencing? Journals remain the single most important source. There is a problem with dissemination and a lack of awareness about repositories. There are also questions about the efficacy of the peer review process. Regarding changes in the bioresearch landscape, we have seen increasing usage analysis linked to large data sets. Open access declarations, recommendations and mandates had a significant impact on the community.

UKPMC has been carried out in partnership with the University of Manchester (which hosts the service, builds "small-scale" developments, engages the HE community, shapes future R&D) and the European Bioinformatics Institute (creates links to the data, integrates it with other repositories, develops discovery interface). The British Library is the prime contractor and manages the grantee database, provides long-term preservation, marks up author submissions, creates marketing collateral, and promotes to the broader user community. UKPMC includes three phases: implementing mirror, marketing, and connection of funding streams to outputs. UKPMC went live on 8 January 2007.

CIBER was contracted to do additional studies on weblog analysis. The data showed the UK to be a significant player in the biomedical area (with a four-to-one ratio of ISI indexed articles where our GDP ratio compared to the US is about six-to-one). Our impact in biomedical research is relatively large. When choosing biomedicine, a problem arises in determining how granular a view should be taken with respect to user information and services provided.

CIBER data on article discovery behaviour contrasted general and biomedical sciences and showed that many researchers are not visiting the library frequently and do not appear to be using specialised A&I services, although they are interested in alerts (Fig. 15). This is an indication that alerting services should be more highly prioritised. Figure 16 shows median length of page views by biomedical researchers in seconds and demonstrates that people are spending very little time looking at online arti-

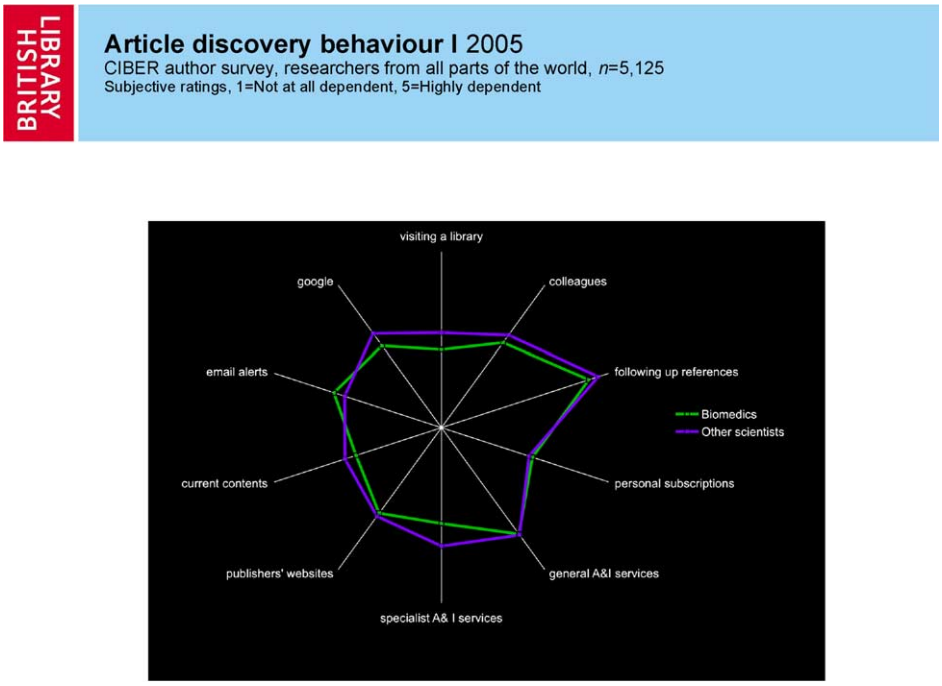


Fig. 15.

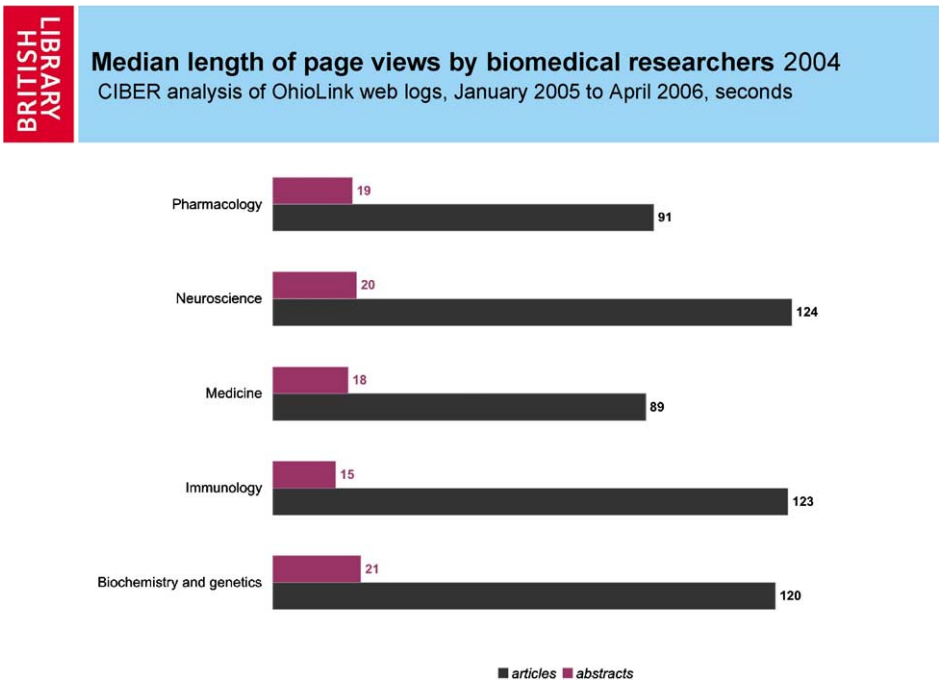


Fig. 16.

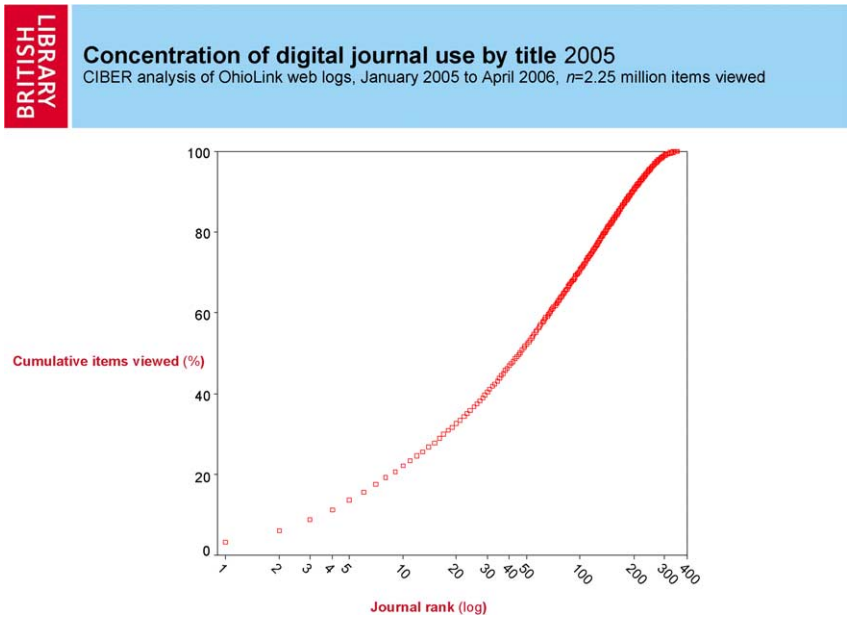


Fig. 17.

cles (thus they must be doing a lot of browsing). This is an area where we need to do more analysis and research. Behaviour is consistent across subdisciplines. Figure 17 ("Concentration of digital journal use

by title") shows the famous long tail problem – higher ranked journals produce the highest number of viewed items. This suggests the value of a comprehensive collection.

In the future we hope to do a range of things around creating a community feel for the UKPMC basic service, enhancing the content, and plugging in bibliographic data. Our intent is to place UKPMC at the centre of UK bioscience research so as to have a central point where researchers can gather. We want to create community-type services based on researchers' feedback and the data gained from various user studies. We also want to create a UKPMC user panel to obtain face-to-face feedback from researchers themselves.

We have carefully examined our overall strategy and come up with some of the same answers as you with respect to what the library should do in the future. The world is going digital and the answer lies in having a better understanding of our users, providing remotely-accessible content, creating communities, delivering material across the Web, and preserving material for future researchers. All of this must be driven by what researchers themselves need to develop their research.

Metrics as a replacement for the RAE

Bahram Bekhradnia, Director, Higher Education Policy Institute (HEPI), Oxford

(I ran the last three research assessment exercises and now run the Higher Education Policy Institute which is entirely objective and evidence driven.)

The Research Assessment Exercises (RAE) is a peer review process that enables selective funding and provides an indicator of comparative (and absolute) quality. In addition, the RAE provides information with international currency and recognition, and management information used within universities. The government's latest proposal for the RAE came, extraordinarily, in the budget and we have yet to receive a good explanation for what is driving this decision. The proposal is to scrap the current system based on peer review – to scrap the panels that do the assessments – and to base future funding on quantitative measures.

Much of the current criticism around the RAE was identified in a speech I gave in 2002. A perennial concern is that long-term research would be penalised by the need to produce outputs that can be judged. This will give rise to indicator chasing which in turn will affect staffing decisions. As a department is judged by the quality of its output, it will focus on recruiting quality people (like a transfer market in football) and will discriminate against younger staff with shorter track records. There were concerns about cost estimated at £37 million in 1996 (up to £100 million in 2001) – including opportunity costs) and damage to non-research activities. All of these concerns were addressed in the guidance and rules given to the panels. While many of these concerns may be valid, it must be noted that that any assessment process of this magnitude (a billion pounds hinges on it) is a competition, and as such will give rise to odd and compliant behaviour. The bottom line is to ensure that the behavioural consequences are benign. It is ludicrous and even counterproductive to suggest that one can have an assessment process that does not affect behaviour.

What did Gordon Brown say was wrong? He cited many of the same concerns, including cost and discouraging of interdisciplinary work. One new point he offered was that it discourages user-focused research. But why did these concerns lead to the current proposal?

I identified two arguments. The first is what the government claims to be prohibitive cost. The government's cost figure for the RAE – including compliance and opportunity costs in the institution – is £45 million. (An unpublished HEFCE internal audit report puts it as high as £100 million.) Even the latter

is less than one percent of the total grants to be allocated. I do not know of a resource allocation system in the world (including the research council system in the UK) whose cost is less than one percent. Thus the cost argument has either not been thought through or is bogus.

The other argument is for moving to metrics, apparently based on a high correlation between Funding Council QR grants and research council grants. But that correlation is a function of size (big universities get more QR and research council grants). For the government to argue the need to eliminate the RAE and move toward metrics based on this strong correlation is both bogus and fatuous. It provides no basis at all.

I am concerned with three issues: scope, process, and substance. With respect to scope I have already noted that the purpose of the RAE is to fund research selectively and to provide an indicator of comparative quality, information with international currency, and management information. The trouble is that it will only provide a basis for funding, not an assessment of quality. It tells you little about a department/university/individual other than the amount of money raised or number of publications produced.

With respect to process, the proposal constitutes one of the poorest pieces of policy making in 30 years. Why the Treasury? Why the rush? Why now? The original 1996 proposal was to scrap the 2008 RAE. The original (and subsequent) proposals were barely thought through. No rationale is provided in the consultation for the proposals. The consultation paper failed to ask any important questions, it simply offered a list of alternative models. The final decision paper was only slightly better (although it *was* better): the humanities have been saved (and will continue to be judged on the basis of peer review). Other subject panels are retained but in an ambiguous role.

Concerns regarding substance are particularly worrying. Quality judgements will be replaced with something quite different. Foreseeable behavioural consequences will include a rush for grants and staffing and equal opportunity issues. If this proposal goes through, the rate of failure for applications in the Research Council grant process (currently 80%) will increase even more, as will the cost. In the current system, quality is based on the whole submission (multiplied by the number of people in a department in which everyone counts to some extent). In a system where all that matters is the money brought in, those who bring in the money will be the stars and the rest will go out the window. As to consequences for the cost of compliance, currently there are multiple sources of judgments (RC and charities on the one hand and the Funding Council on the other). That will go. If there is a direct relationship between QR money obtained and the amount of money from RC and industry, then we are effectively looking at a subsidy for industry research; it will require researchers and universities to offer cut-price research to industry. If all research money is dependent on getting grants and contracts, it will kill curiosity-driven and unpopular research. The future will not be better than the past.

What are the alternatives to what currently exists? If we want a process that produces judgments of comparative quality, peer review must be at its heart. However, metrics can play a part and can serve as a trigger for peer review. I predict that citations will prove unworkable as a mechanism to distribute funds and that the science subjects will look at the humanities, who have been spared this new regime, and ask why they too cannot have a rigourous assessment mechanism and why they must be subject to these metrics.

Funding agencies and their needs

Dr Liz Allen, Senior Policy Adviser, The Wellcome Trust

I will discuss some of the perspectives we face as a funding agency in terms bibliometric and bibliographic information use. As one of the world's largest biomedical funding charities (second only to the

Bill and Melinda Gates Foundation) our goal is to *foster and promote research with the aim of improving human and animal health*. We have a broad remit that covers biomedicine – from pure basic biomedical research to history of medicine research, humanities, and the arts. We currently fund some 6,000 researchers in 50 countries at an expenditure of about £500 million annually. The Wellcome Trust has a history and ongoing interest in publication analyses and contributes to over 4,000 research publications yearly. We are committed to open access publishing and data sharing.

We currently use a range of online databases including Scopus and PubMed, and will soon be using PubMed Central. Current projects include an annual overview of publication output associated with Wellcome Trust, initiative-based reviews and evaluation (with cohort tracking as a subset), monitoring of a deposition into UKPMC, and bespoke research (including into our own policies). We are looking at whether publication output is associated with different funding types and which ones produce greater output in terms of volume, citations, and where they are published. With respect to individual papers, we want to know who collaborated on and funded them. With respect to journals we want to know if they are open access. We are interested in self-citation behaviour and in quality indicators (e.g. F1000, ESI, WoS) and are using databases to help validate that.

Much of our work is dependent on the nonresearch text in research papers including address fields, acknowledgements section, and additional information at the bottom. Currently this is data that is not very searchable. There are a range of databases that do not really speak to each other which makes the process complex.

Figure 18 takes as an example a paper published in *Science* to demonstrate how we have used four databases – PubMed, Scopus, Web of Science, and Faculty of 1000 – without having access to the full text. PubMed provides full reference details, indicates the association with The Wellcome Trust and provides the lead author's address. However, it does not provide addresses for other authors (an issue for

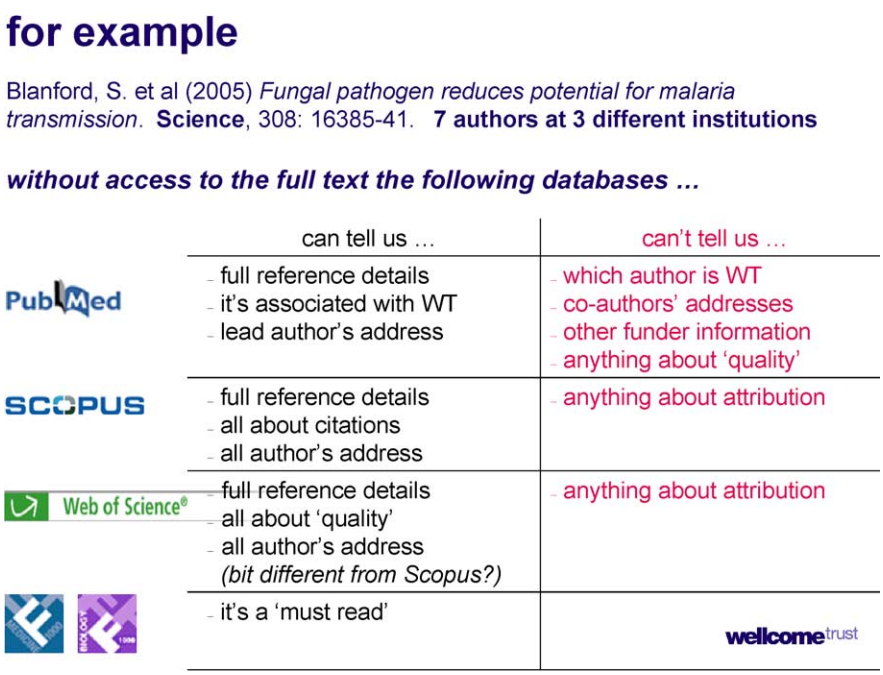


Fig. 18.

us). Scopus, which is used to complement PubMed, provides all authors' details, references, and citation data but does not provide information about attribution or funding sources. The same issues apply to Web of Science (which inexplicably offered up different author names than Scopus). Faculty of 1000 – which consists of groups of renowned academics who recommend papers and place them on their website – is used as a validation exercise.

A major issue concerns variation in acknowledgement practice and convention where we are faced with time-consuming manual exercises. An absence of benchmarks makes long-term comparison difficult. On the positive side there has been a "revolution" in access to publication outputs permitting us to reach data more quickly. Many databases offer a range of tools and functionality, although there is concern with variations in compatibility and export functionality. It can also be difficult to use data in a complementary way. Text mining and searching functionality have improved – on PubMed and UKPMC it is possible to search the acknowledgement section.

Understanding user behaviour and its metrics: what is missing?

David Worlock, Research Fellow, Outsell Inc./EPS Ltd.

A few reflections: We are in danger of taking a slightly complacent view. The power of users in networks seems to be much talked about but totally under-recognised by those who source, publish, assess, collate and curate information and serve it back to the originating user. The user is powerful and changing. I am not sure that any of the things we do with what we choose to call user behaviour (and I want to quarrel with what user behaviour means) or indeed what we choose to call metrics, actually measures or evaluates something fundamental about change in our society.

There are publishers, such as Tim O'Reilly, who are capable of talking change. O'Reilly (with whom I was recently on a conference platform) talks and acts change because he reinvents what he does out of his users' experience. There is an important issue here that is not about usage but rather about behaviour expressed differently.

In the "new" science publishing paradigm (Fig. 19), primary content (the base of the triangle) is increasingly commoditised. We speak of this data (from open archives, repositories, publishers) as though it were some great currency even where the rate of exchange has changed so radically that currency is valueless. What would happen to the world we are describing if users decided that the fundamental material they wanted was the evidence that is produced by research work? What would happen if communications at conferences were at least as valuable and important as communications written as articles? I think the journal – with the exception of a few big branded ones – has a dangerous totter about it. I foresee the development of branded databases containing articles (with many of those articles connected to evidential resources) and the informal communication of the STM community becoming at least as important as the formal.

We mistakenly believe that our STM world is immune from societal influence. In fact it is under huge pressures of its own, including those of the e-science movement and of increasing communication between remotely-cited researchers (communications that cannot be monitored or validated which may indeed have a huge role to play in how reporting takes place). We have reached a value triangle that includes A&I and mapping individualisation and we are only beginning to move into what O'Reilly would call Web 2.0 but what I would call community services.

We are all concerned with the nature, depth and connectivity of communities. And communication in those communities may itself be a very valuable part of this whole activity. We need to move towards

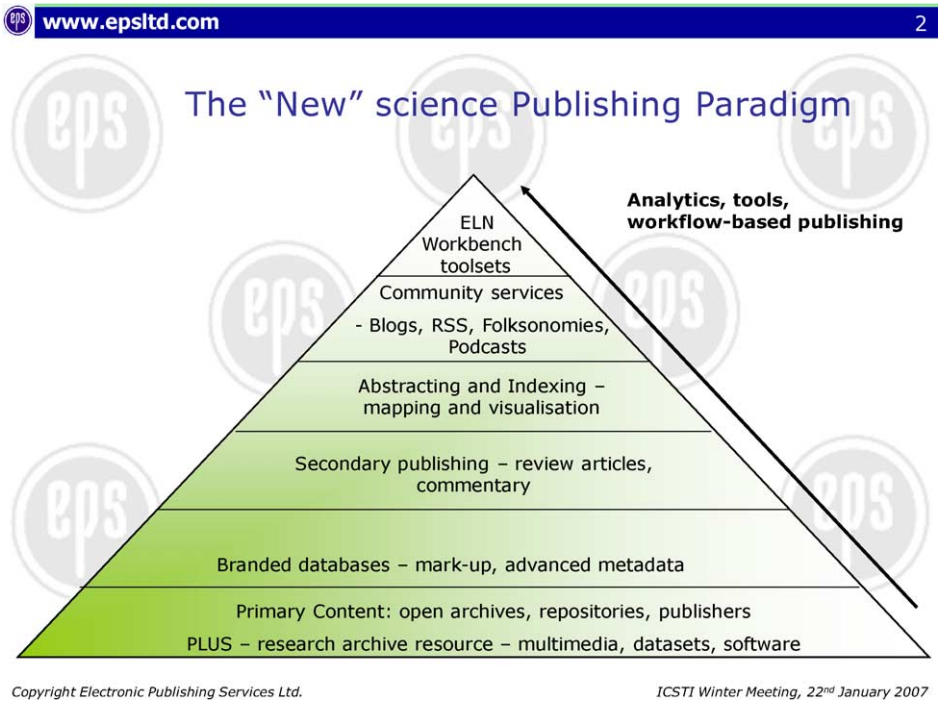


Fig. 19.

what I call the process of this diagram, adopting some of the current language including ELN, work bench, and tool set – the dashboard of the researcher in which his content is going to be plugged. But do we understand what needs to be on the dashboard? What needs to be in the ELN? Do we understand the ethical and compliance issues?

Last year EPS carried out a project to evaluate current research in six key areas. We found very poor data about buyers of scholarly journals and noted that existing market studies lacked comprehensive coverage. We also saw that more survey data provided directly from the publishers themselves (anonymised) would help fill the gaps. We looked at a wide range of estimation on the supply side and the evidence here was poor. We looked at journal usage where we found consensus on the incompleteness of the picture. These are based on user surveys which, as David Nicholas noted earlier, can be misleading as survey answers are not always honest. We found transaction logs to be a very valuable form of analysis, but hard to get at. Evidence regarding the value of citations and impact factors was inconsistent. The current marketplace tends to favour spot surveys over long time frames which makes it difficult to obtain real time series. There is positive evidence of the cost impact of alternative methods.

One major gap was around behaviour. I would define behaviour very differently from Hugh Look's definition earlier today. What he has done is legitimate, but in most Web-penetrated marketplaces we are returning to the idea that the user himself has had his behaviour moderated by the facts of the Web. What we do not know in STM is how researchers now work. We assume that researchers work the way they always have, that research methodologies have not been impacted and the Web just exists. I would argue that this is not so.

I do not know of any community using the Web whose behaviour has not been moderated by the form of its access to information. And that is the case here. The nature of research being done is therefore

changing. How can we begin to design the dashboard without knowing about behaviour modification? How can we know whether the publishing value models that we are playing with fit the evolving behavioural model of how research is accomplished?

What are we doing? We are going into the past to look at the usage patterns of the past, at the way in which things have traditionally been done. We need to track the changing behaviour of a networked research world going into the future. Only if we have our fingertips on this in our libraries and publishing houses, will we actually be able to respond to those pressures.

Closing remarks

Bernard Dumouchel, Director-General, CISTI/NRC Canada

Thanks to our speakers we have established our objectives. We may not have achieved all the solutions, but we have highlighted the challenges. We have heard about current practices and listened to evidence of what we will use now and in the future to build new, different, and adapted information services in the broadest sense – as publishers, authors, librarians, and members of society. This public conference has delivered a number of key messages about user behaviour, how we measure information obtained on user behaviour, and how we implement that into our information services. We have seen that impact is crucial in an era where outcomes are key to our ability to continually offer value-added information services. We also heard from our funders about what they expect from us. We must come out of this public conference acknowledging all of these elements and looking at ourselves, our services, and our activities. We must move forward by looking at how this changing paradigm will add relevance to our activities and help us to face the challenges ahead. I want to thank all the speakers for attending this public conference. I will close by reminding you that we will pursue these issues in a similar vein, but with greater focus on SciTech and information *per se* at our conference in Nancy (France) on June 21st 2007. I hope to see you there. Thank you.

Acknowledgements

Special thanks to Penny Bunch, Emma Cass, Elisabeth Maître-Allain and to Antoine Raulin for their efforts in organizing the workshop, and to Morgan Wolfe for her contributions to the summary proceedings.