If we do not understand our users, we will certainly fail

David Nicholas and Paul Huntington

CIBER

UCL Centre for Publishing

School of Library, Archive and Information Studies

www.publishing.ucl.ac.uk

We might be failing

- **The digital transition.** As our users have migrated to the Web in huge numbers we know less about them than we ever did.
- **Marginalised and anonymous.** As users increasingly access services from their office, home and Google we have become marginalised, not even credited for the digital services we deliver.
- **Huge consumer choice and disintermediation.** Add to the woes. Need to know what choices made and why

But there is hope in the digital environment

- Users leave fingerprints.. Need to monitor users like never before and have the ability to monitor them like never before.
- Deep log analysis helps make sense of the fingerprints: from hits to users, to outcomes (ScienceDirect, OhioLINK)
- Developed evidence base on behaviour of virtual scholars
- Pick out the key findings, which should be of value to Librarian's
- They concern: 1) popularity, 2) user diversity, 3) bouncing, 4) online reading, 5) digital visibility, 6) search engines, 7) trust
- Examples come from investigations of ScienceDirect, Synergy, EmeraldInsight, OUP Open and OhioLINK

Popularity of digital services

- Users phenomenally active, increasingly interested in e-journals (and we about to look at e-books, learning material)
- Synergy: more than 500,000 people visited a month, recording nearly 5 million views
- OhioLINK: 6000 journals and all bar 5 not used within month
- Nucleic Acids Research (NAR): 17,150 downloads made in a single month and usage increased by 150% in two and a half years
- And its not just subscribers - EmeraldInsight: two-thirds of visitors non-subscribers and NAR's in the majority
- Scholarly publications in demand and demand driven ever upwards by improved access – Big Deals, search engines, OA

User diversity

- Must move away from hits to users. Very real differences between various types of user, especially in regard to their subject field; academic status and geographical location. We have also examined - and found in some cases big differences - according to gender, type of organisation worked for, type of university, attitudes towards scholarly communication
- One size does not fit all. Need to drill down.

Subject diversity – some examples from ScienceDirect

- **Age of material.** Users from Economics (71%), Engineering (71%), the Social Sciences (69%) and Computer Science (70%) made most use of current (one-year old material); Material Science (51%) and Mathematics (52%) least.
- **Number of journals consulted in a search.** Users from Material science (39%) an Mathematics (38%) most likely to view 2 or more journals in a search. Those from Medicine (69%) and Computer science (69%) most likely to view just one.
- **Return visits.** Computer Science (80%) and Physics (79%) recorded high percentages of repeat visits over a five month period and Engineering the least (46%)

Other user diversity

- **Age and abstracts.** 14% of those aged 36-45 just undertook an abstract only session; double that for those aged over 56.
- **Gender and PDFs.** Men more likely to undertake a PDF only session (37%); women (22%)
- **Geographical location and 'productive' searching.** E. Europeans (47%) recorded a high percentage of searches resulting in zero returns. North Americans very 'successful' searchers - 74% of their searches resulted in one or more matches. Overall though, Germans the most 'successful' searchers (more hits, less zero searches)

'Bouncers'

- Typically 50-60% viewed no more than three items in a visit and then left;

- Significant number did not return - within a year 50-60% did not come back. For some, user loyalty at a premium;

- Search a variety of sites to find what they want;

- Suggesting at **best** a checking-comparing, dipping sort of behaviour that is a result of search engines, a shortage of time, gateways and huge digital choice. *Flicking.* At **worse**, massive failure at the terminal? Takes us to **outcomes**.

Online reading, outcomes and digital literacy

- Bouncing raises questions about outcomes, as does another possible 'dumbing down' characteristic – online 'reading';

- 1) People spend more time reading shorter articles online. 2) As the length of a paper increases the greater the likelihood that it was viewed as an abstract and there was less likelihood that the item will be viewed in full text. 3) What of downloads, the outcome/satisfaction 'gold' standard? Digital osmosis.

Digital visibility

- Improved access and increased visibility leads to increased exposure and use of older scientific material – have we woken up to this? Search engine searching a key here (following slide)

- ScienceDirect downloads to material older than 5 years in case of Materials Science (59%) and Physiology (64%)

- Increasing visibility – sales, home pages etc (following slide). Big lessons for libraries here.

'Sales' help too
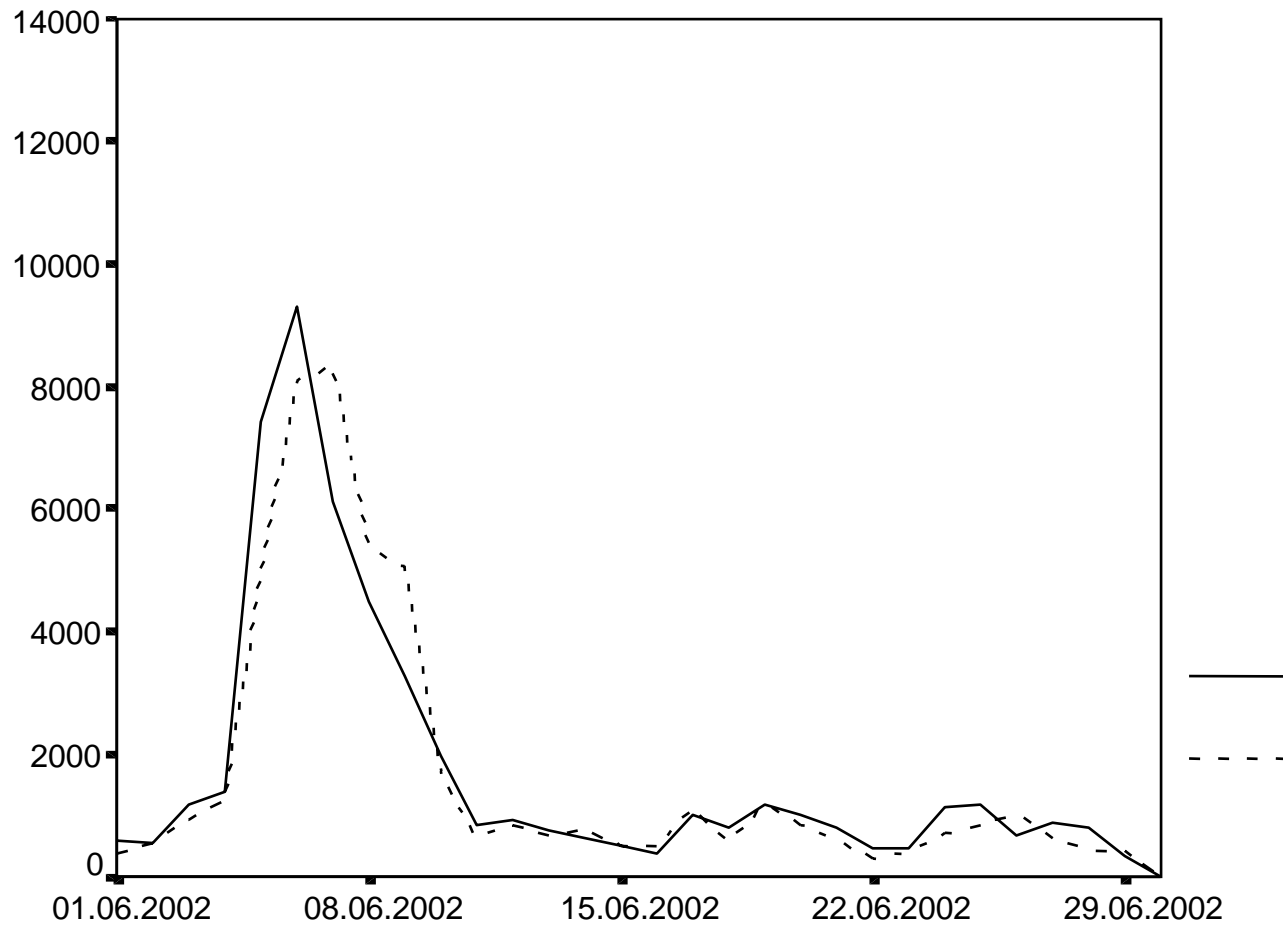
Search engines

- Not surprisingly the method of navigation/searching/browsing adopted determines what is viewed and used
- People using search engine were:
  - far more likely to conduct a search that included a view to an older article;
  - more likely to view more subject areas, more journal titles, and also more articles and abstracts too.
  - more likely to be 'bouncers'
- Undergrads most likely to use search engines: 46% had compared to 26% of postgrads, 19% of researchers, 15% of professors.

Trust and brand

- Complex and difficult to ascribe
- Determining responsibility a problem and a surprise – Tesco health example is illustrative
- Authority and relevance to be won (and cross-checked).
- Differences between age groups
- Special problems for libraries: 1) – searching from home, via Google to Oxford Open – no Library involvement; 2) searching from office to EmeraldInsight – no Library involvement obvious;

Kicking user studies up a level- analysis examining age of material consulted

- Examined the relationship between responses to the question 'It is NOT important to have access to research articles that were published more than 10 years ago' to the age of the articles viewed. As might have been expected those who strongly agreed with the statement were most likely to engage in a session where just current material was viewed and those that disagreed were most likely to engage in a session where only older material was viewed. Powerful triangulation here.

Conclusions and implications

- Choice and common & multi-function retrieval platform changing everything; making us behave as consumers and should question assumptions about today's scholar
- Horizontal rather then vertical searching. Dumbing down – life critical services?
- DLA raises questions and provides 'grip'. When people are provided with a digital service things don't go as expected.
- Need to get closer to user but maybe moving further apart.
- Time to turn away from hits and move to users and outcomes.
- Future. Watch that OhioLINK data (Charleston) Planning to work with major university libraries to demonstrate what can be achieved. Also researching new publishing models (UKPMC).

# Some references
(see also www.publishing.ucl.ac.uk)

1. Nicholas, D., Huntington, P. and Watkinson, A. Scholarly journal usage: the results of deep log analysis. *Journal of Documentation,* 61(2), 2005, 246-280.

2. Nicholas, D., Huntington, P., Dobrowolski, T., Rowlands, I., Jamali, H. R. & Polydoratou, P. Revisiting 'obsolescence' and journal article 'decay' through usage data: an analysis of digital journal use by year of publication, *Information processing and Management*, 41(6), 2005, 1441-1461.

3. Nicholas D, Huntington P, Williams P and Dobrowolski T. 'The Digital Information Consumer in *New directions in human information behaviour*.' Edited by A Spink and C Cole. Kluwer Academic, 2005

4. Nicholas D and Huntington P. Digital journals: are they really used? *Interlending and Document Supply*, 34(2), June 2006, 74-77

5. Nicholas D, Huntington P, Jamali HR and Tenopir, C. Finding information in (very large) digital libraries: a deep log approach to determining differences in use according to method of access. *Journal of Academic Librarianship,* 32 (2), March 2006, 119-126

6. Nicholas D, Huntington P, Jamali HR and Tenopir C. OhioLINK – ten years on: what deep log analysis tells us about the impact of Big Deals. *Journal of Documentation,* 62 (4) July 2006, 482-508
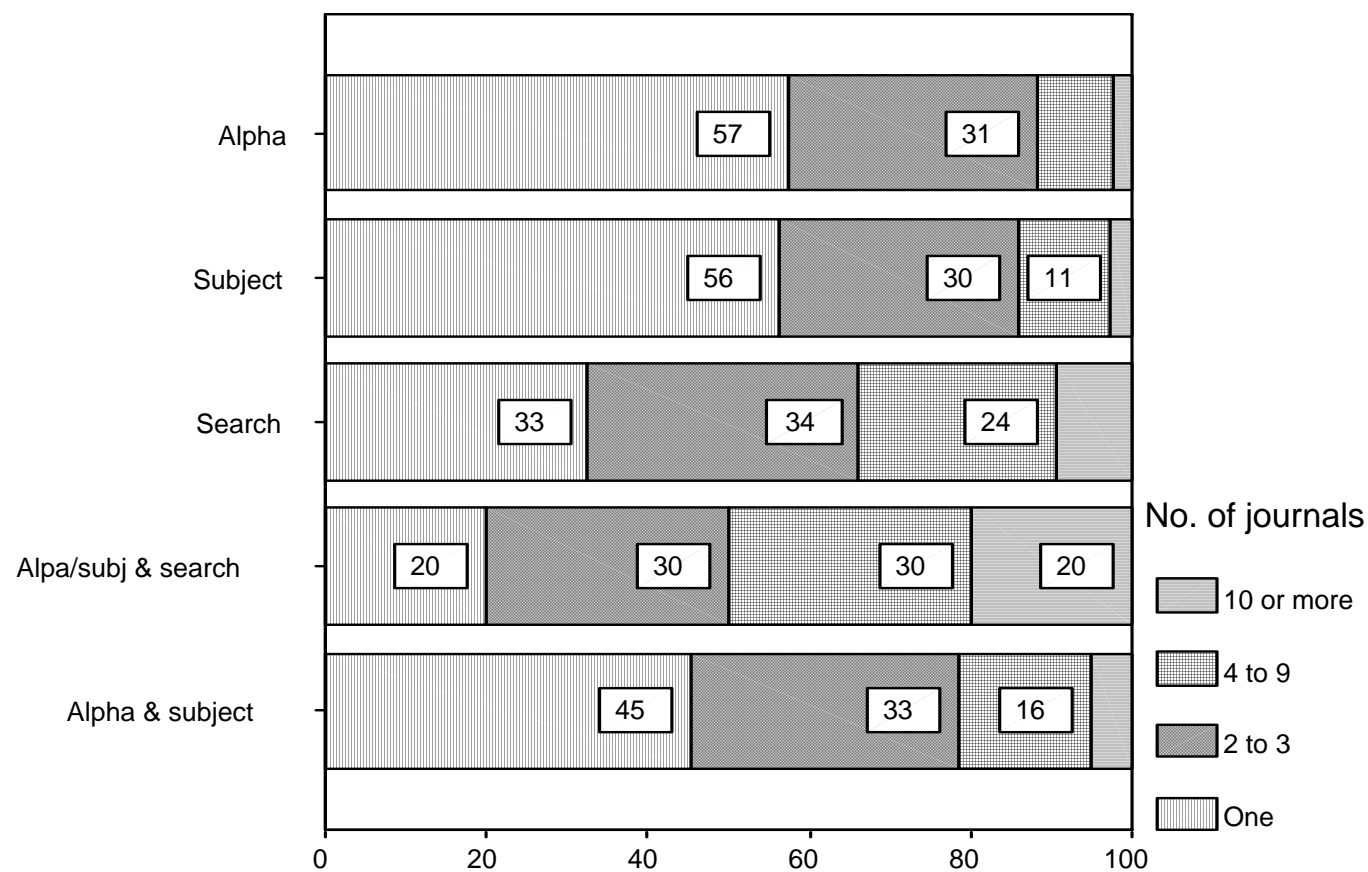
Appendix

NAR: monthly articles viewed by referrer
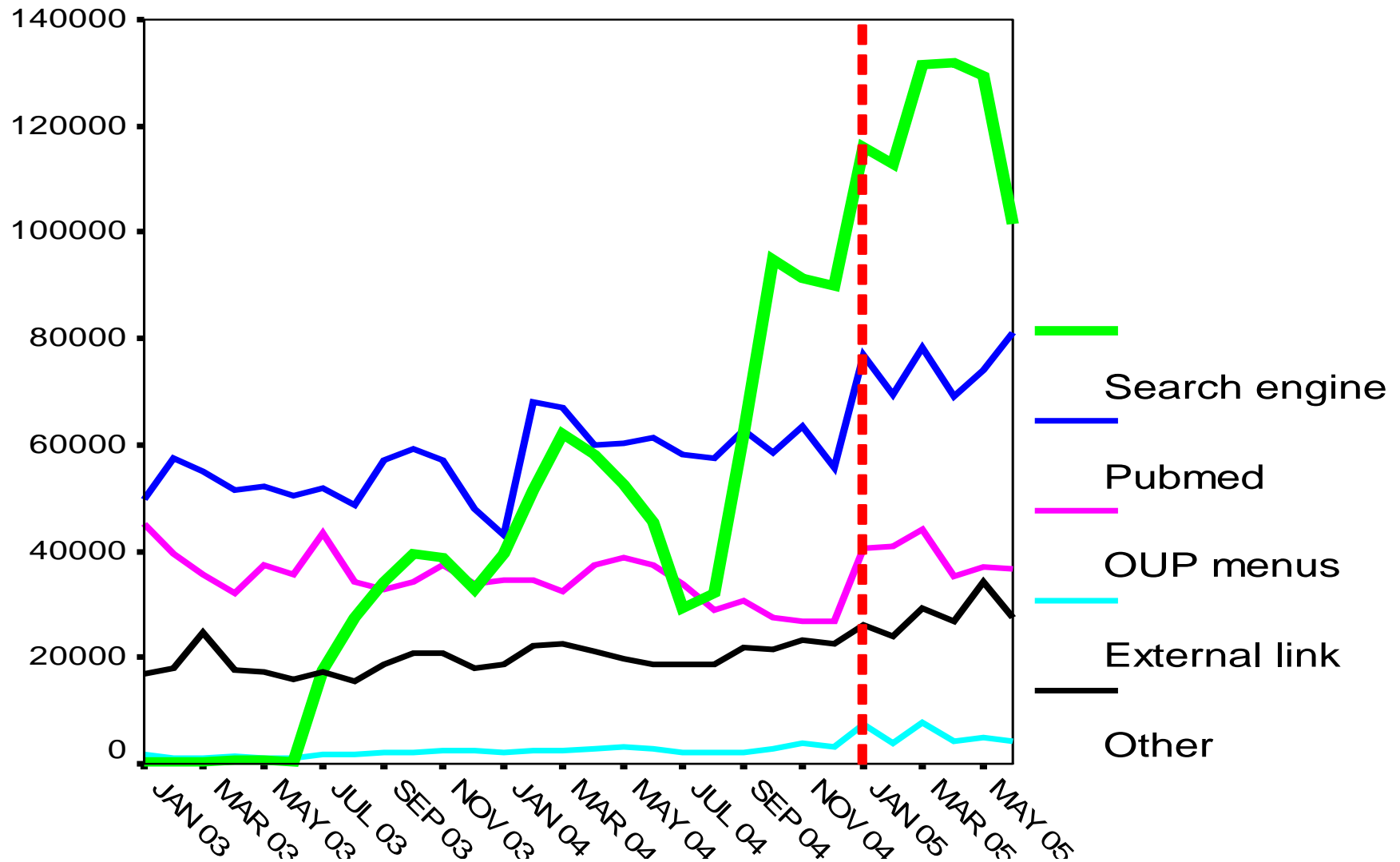
Daily article views for two & half years
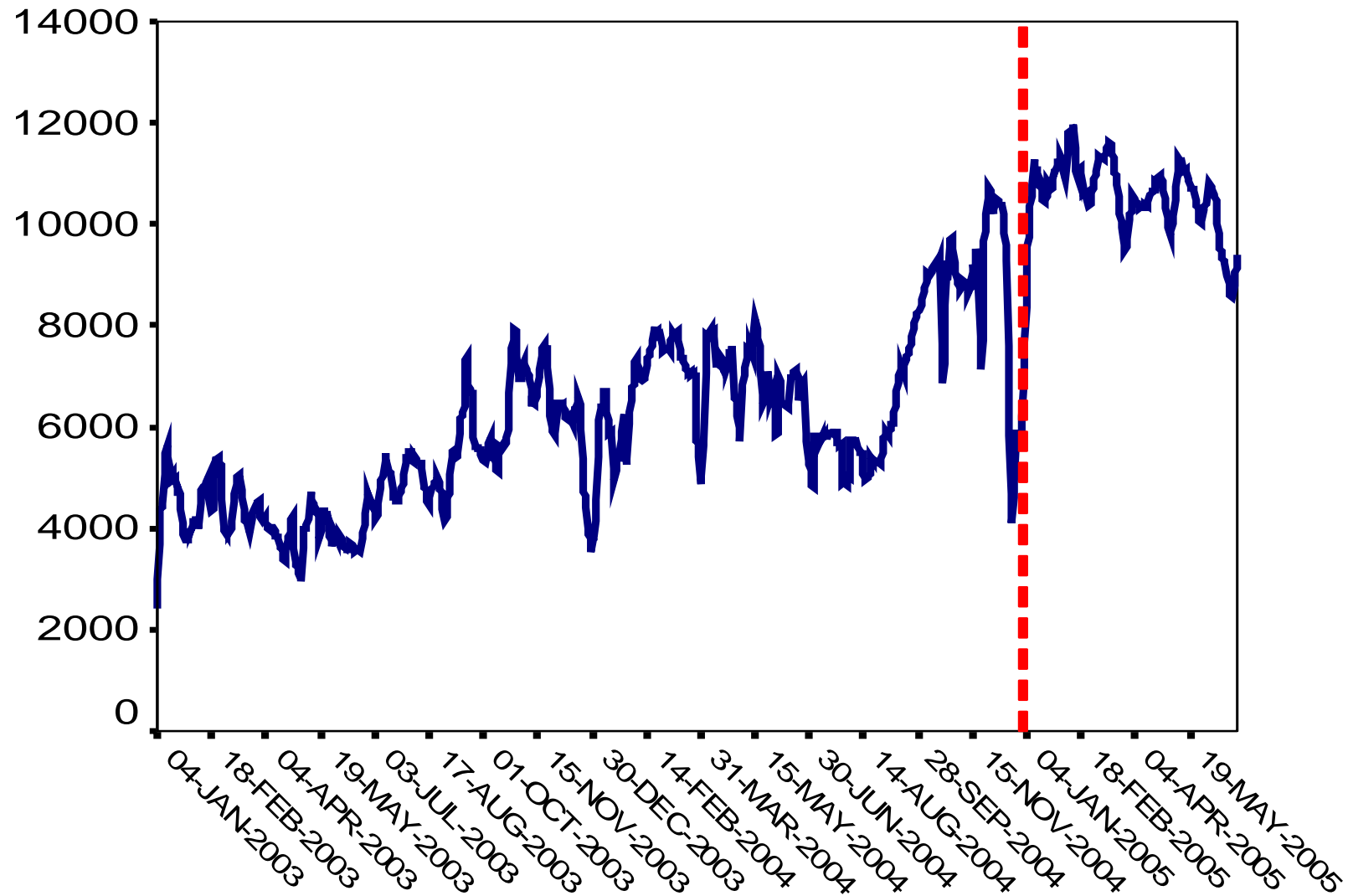
Number of different journals viewed by access method (OhioLINK)

Subject of journal by date of material viewed (OhioLINK)