What researchers want

Michael Jubb
Research Information Network

Decoding the Digital
27 July 2010
Some propositions

- the volume of research undertaken worldwide has increased, is increasing, and will continue to increase
  - and more of it will be done collaboratively
- researchers are both producers and consumers of research outputs
  - but they don’t necessarily share the same interests
- Governments invest in research because they believe it has a positive impact on society and the economy
  - and they want to maximise that impact
- the costs of research, and of higher education, have increased, are increasing (and ought to be diminished?)
  - cost-effectiveness an increasingly-dominant theme in current economic climate
The Role of Information in Research:

a crude model

defining a set of research questions, issues or problems
identifying relevant existing knowledge
accessing, analysing, and evaluating existing knowledge and data
designing a methodology for generating new knowledge
applying the methodology and discovering new knowledge
combining old and new knowledge to answer research questions and to enhance understanding
disseminating the outcomes of research in a form that is both sustainable and retrievable
The Research Process: Animal Genetics
The Research Process: Transgenesis and Embryology
The Research Process:
Epidemiology
The Research Process

- differs even in apparently similar areas of work, and also between teams.........
Composition of Research Groups

big science vs small science
- small teams typical in life and physical sciences
- amorphous and overlapping associations with other teams
- “primary research engagements tend to be local”

divisions of expertise, labour and knowledge exchange
- PI/leader, senior researchers/lecturers, associates, computational specialists, postdocs, PhDs, technicians........
- dangers of surveys that look at individual responses divorced from context
### Content: what do researchers want to find and use?

<table>
<thead>
<tr>
<th>Research Resources</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Journal articles</td>
<td>99.5%</td>
<td>0.5%</td>
</tr>
<tr>
<td>Chapters in multi-authored books</td>
<td>97.0%</td>
<td>3.0%</td>
</tr>
<tr>
<td>Organization’s web sites</td>
<td>90.8%</td>
<td>9.2%</td>
</tr>
<tr>
<td>Expertise of individuals</td>
<td>90.1%</td>
<td>9.9%</td>
</tr>
<tr>
<td>Conference proceedings</td>
<td>85.8%</td>
<td>14.2%</td>
</tr>
<tr>
<td>Monographs</td>
<td>83.3%</td>
<td>16.7%</td>
</tr>
<tr>
<td>Datasets – published or unpublished</td>
<td>62.0%</td>
<td>38.0%</td>
</tr>
<tr>
<td>Original text sources, e.g. newspapers, historical records</td>
<td>61.5%</td>
<td>38.5%</td>
</tr>
<tr>
<td>Preprints</td>
<td>54.7%</td>
<td>45.3%</td>
</tr>
<tr>
<td>Non-text sources, e.g. images, audio, artifacts</td>
<td>47.0%</td>
<td>53.0%</td>
</tr>
<tr>
<td>Other</td>
<td>18.0%</td>
<td>82.0%</td>
</tr>
</tbody>
</table>
Content: user expectations and needs

- published and non-”published”
  - grey literature, reports, working papers
  - data: raw or refined? mine or yours?
  - websites, blogs, wikis, emails

- quality-assured and non-”quality-assured”? 
  - the good-enough source and/or version?
  - pre or post-publication peer review?

- digital and non-digital
  - perdurance of the book?
  - role of digitisation
E-journals are used

- 98% of titles used in 10 institutions over 4 month period
- over HE sector as a whole, annual growth rate (CAGR) of 21.7%
Researchers in different disciplines behave differently

<table>
<thead>
<tr>
<th></th>
<th>Journal titles viewed</th>
<th>Most popular 5% of journals accounted for % use</th>
<th>Page views (average per session)</th>
<th>Abstract views (% sessions)</th>
<th>Gateways (% page views arriving via gateways)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemistry</td>
<td>196</td>
<td>39.5</td>
<td>3.2</td>
<td>23.3</td>
<td>49.2</td>
</tr>
<tr>
<td>Environmental sciences</td>
<td>248</td>
<td>29.6</td>
<td>3.6</td>
<td>22.7</td>
<td>41.4</td>
</tr>
<tr>
<td>Economics</td>
<td>132</td>
<td>46.9</td>
<td>3.8</td>
<td>30.4</td>
<td>19</td>
</tr>
<tr>
<td>Life sciences</td>
<td>531</td>
<td>38.1</td>
<td>2.0</td>
<td>19.5</td>
<td>65.9</td>
</tr>
<tr>
<td>Physics</td>
<td>204</td>
<td>26.6</td>
<td>2.5</td>
<td>20.1</td>
<td>57.8</td>
</tr>
</tbody>
</table>
But *per capita* usage varies.
Profile of journals varies too

<table>
<thead>
<tr>
<th>Case study</th>
<th>Average impact factor of journals viewed</th>
<th>Relative impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aberdeen</td>
<td>3.0</td>
<td>1.2</td>
</tr>
<tr>
<td>Bangor</td>
<td>2.3</td>
<td>0.9</td>
</tr>
<tr>
<td>Cambridge</td>
<td>5.0</td>
<td>2.0</td>
</tr>
<tr>
<td>Centre for Ecology and Hydrology</td>
<td>2.6</td>
<td>1.0</td>
</tr>
<tr>
<td>Edinburgh</td>
<td>3.7</td>
<td>1.5</td>
</tr>
<tr>
<td>Manchester</td>
<td>3.9</td>
<td>1.6</td>
</tr>
<tr>
<td>Rothamsted</td>
<td>2.6</td>
<td>1.0</td>
</tr>
<tr>
<td>Strathclyde</td>
<td>2.7</td>
<td>1.1</td>
</tr>
<tr>
<td>Swansea</td>
<td>2.5</td>
<td>1.0</td>
</tr>
<tr>
<td>UCL</td>
<td>4.1</td>
<td>1.7</td>
</tr>
</tbody>
</table>
Services: user expectations and needs

- researchers as creators
  - quality assurance and enhancement
  - distribution and marketing

- researchers and others as consumers
  - quality assurance
  - search and navigation services
  - access, 24x7 and permanent
  - links and interoperability
  - text mining (published text as data)

- funders and research institutions
  - assessment and evaluation services
Services: usability and sustainability

- search, navigation and access
  - invigorating competition or wasteful duplication?
- levels of usage of services provided by publishers and libraries
Skills, expertise and competences: user expectations and needs

- specialist research skills and specialist information skills
- what’s easy, and what’s not
  - and how that changes
- ‘information literacy’ approaches and their limitations
- need for enhanced information professional skills in some areas
  - eg management and communication skills; bibliometrics
Skills, expertise and competences:

- Specialist advisor in copyright/IPR
- IT support expert
- Manager for metadata
- Technol spec't (e-access)
- Technology specialist (VREs)
- Facilitator for e-learning
- Custodian archives, spec collections
- Manager of e-science/grid datasets

Graph showing the skills distribution among different roles and groups.

Legend:
- Library directors
- Library staff
- Researchers
Some conclusions: users as creators and consumers

Understanding e-journal investment, use and research outcomes

Conclusions in summary

This diagram summarises the previous discussion and enables us to get more a sense of directionality.

There appear to be three main drivers. The first is that investment drives use (actually it is a precondition for it). The reverse argument, that use drives library spending is weak, probably because of the bundled nature of journal purchasing.

Direct linkages between library investment and research success are weak, but then the two things being measured: a spending decision on journals and subsequent research performance metrics are a long way from one another in time as well as conceptually.

The summary model suggests a positive feedback loop between the use of e-journals and research outcomes. This is completely intuitive. The beauty of this model is that it shows that use is a strong predictor of research success and the biggest factor in the model by some distance.

Limitations
This is a small-scale study to explore the concepts behind path modelling to see if they could be used in studies of library return on investment. But it looks promising...
Questions?

Michael Jubb

www.rin.ac.uk