



Creating, using and communicating knowledge: a researcher perspective

Michael Jubb

Research Information Network

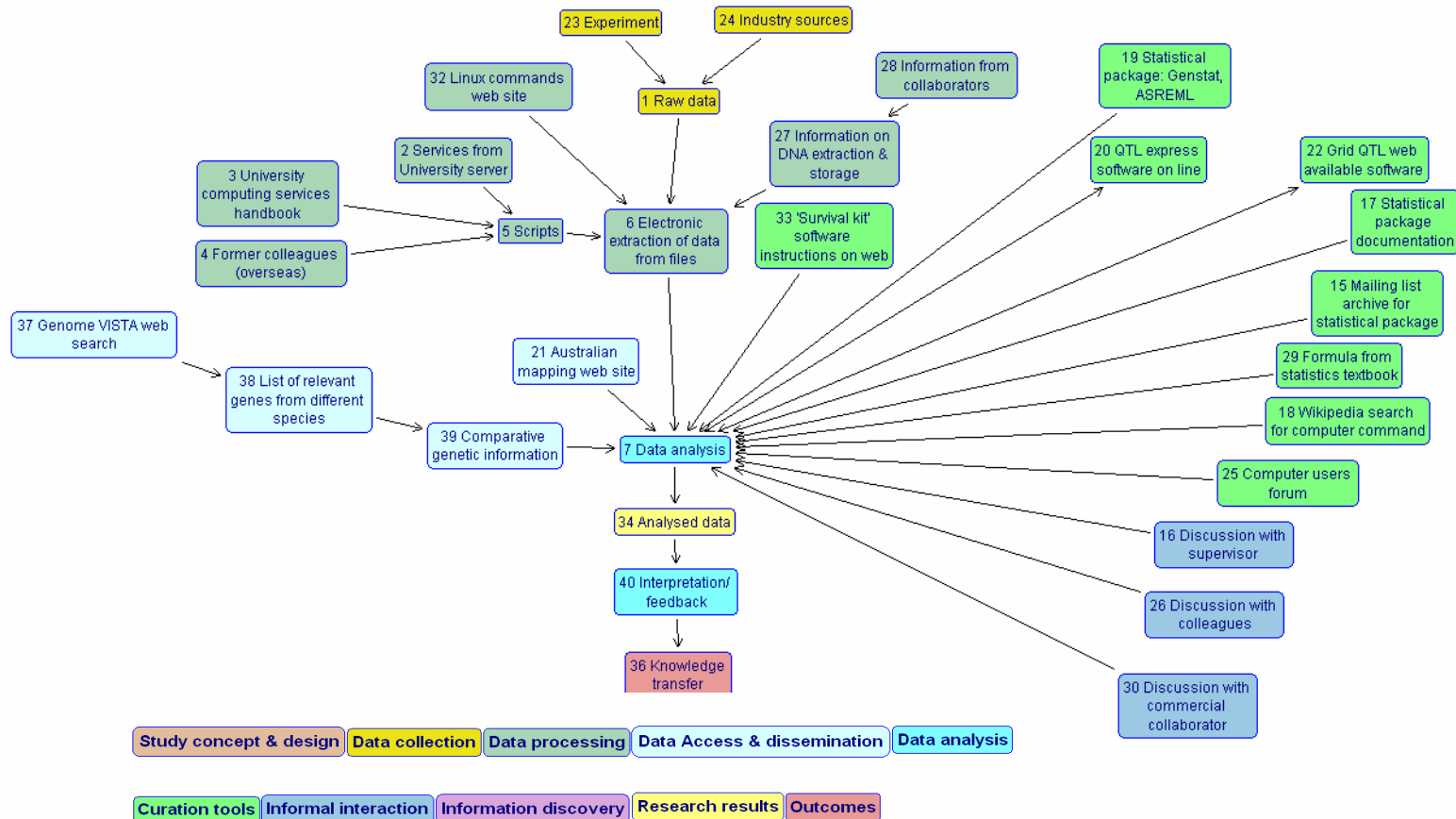
APE 2010 Conference

Berlin, 19-20 January 2010

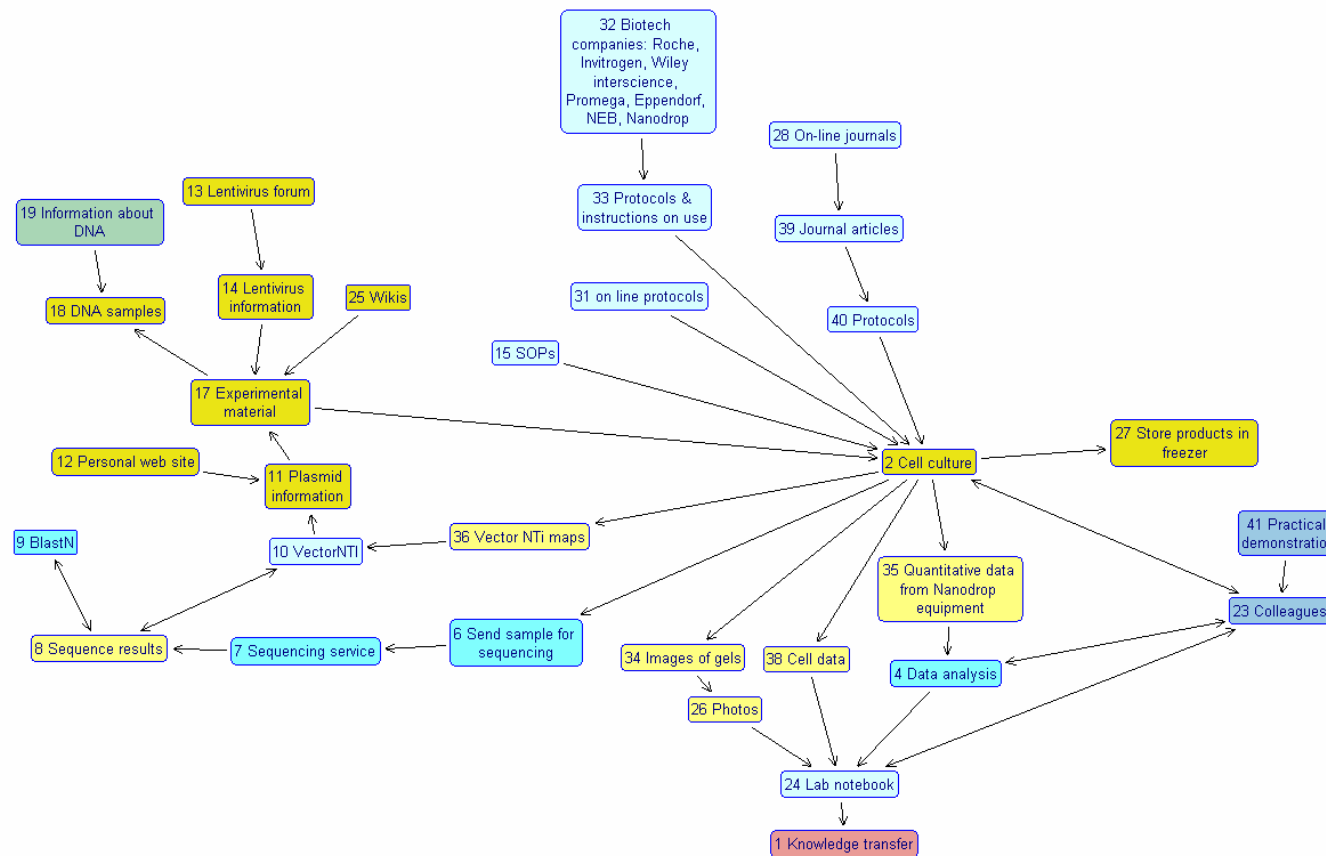
Information in the research process: some verbs in the lifecycle

- gather
- evaluate
- create
- analyse
- manage
- transform
- present
- communicate
- disseminate

The research process: animal genetics



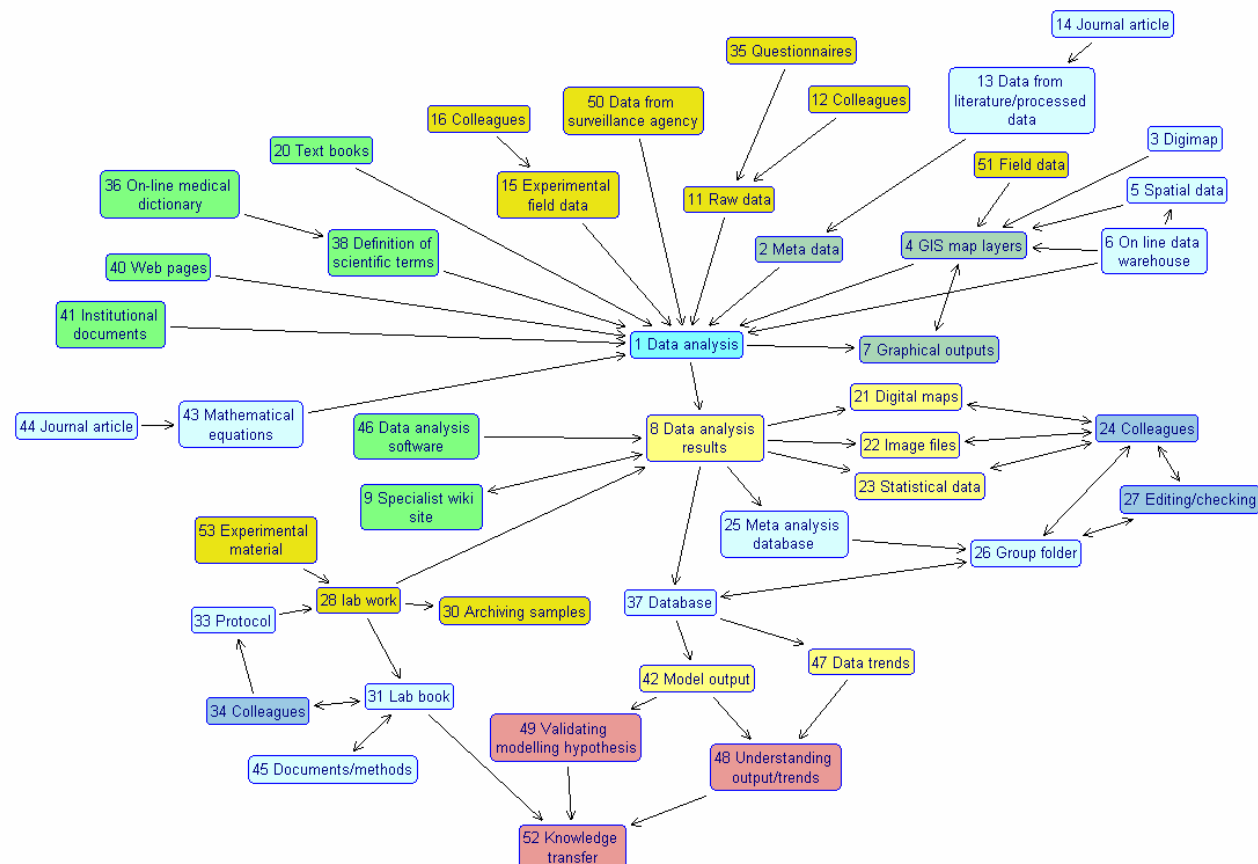
research process: transgenesis & embryology



Study concept & design Data collection Data processing Data Access & dissemination Data analysis

Curation tools Informal interaction Information discovery Research results Outcomes

research process: epidemiology



Study concept & design | Data collection | Data processing | Data Access & dissemination | Data analysis

Curation tools | Informal interaction | Information discovery | Research results | Outcomes

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 more typical than often assumed
 - amorphous and overlapping associations with other teams
 - “primary research engagements tend to be local”
- **divisions of expertise, labour and knowledge exchange**
 - principal investigator/leader, senior researchers/lecturers, associates, computational specialists, postdocs, PhDs, technicians.....
 - dangers of surveys that look at individual responses divorced from context

Researchers and the skills gap

- *“...a lot of scientists don’t get information and structures at all. It’s not what they’re trained to think about.”*
- *“.....the idea of quality, provenance, and metadata about data is woefully inadequate in most science training.”*
- *“.....it’s not just about creating the infrastructure, it’s actually about creating the demand for the infrastructure, because.....there’s still a lack of demand.”*
- engagement between researchers and information professionals
 - *“.....now we manage our data, whereas before we didn’t”*
- scalability in training and support, and in promoting cultural change

Discovery and access : some generalisations

- ▣ **Google**
 - ▣ lack of concern about limitations
- ▣ **range of other sites and databases**
 - ▣ limited awareness of what is available
 - ▣ limited time and “learning costs”
 - ▣ find a service you like, and stay with it
 - ▣ importance of (very) domain-specific and (highly) specialist services
 - ▣ “informal discussion” a key source of information and advice
 - ▣ relatively little use of blogs, wikis etc
- ▣ **barriers to access to full text are still a concern**
 - ▣ resistance to requirement to pay
- ▣ **multiple platforms an inhibitor to take-up and use**
 - ▣ even Grid users want to work simply on the desktop

Different roles and activities: who or where is your information coming from?

team leader	senior researcher	senior lecturer	postdoctoral researcher	PhD student
imaging lab group meeting researchers verbal report on studies grants reports lecture on specialist topic journal articles emails phone calls discussions with researchers publishers letters clinics clinical meetings discussion with clinical groups (internal and external)	interviews with psychiatrists handwritten notes from interviews server (image lab) digital observations image files wiki (statistics) manuals online dictionary and wikipedia journal articles pubmed– papers biomed central home pages and websites (institutions) drafts of sci paper peer reviews e-mail (personal and team) my head – 8 years experience	imaging data, department server and online resources data from: imaging resources (fmri data) from wgh [hospital] animal imaging data (qmri) talks/presentation (wgh) internet accessed journal articles (2)	clinical cohort scan database discussions with team leader/ PhD student / clinical researcher/postdocs /prof clinical journal group discussion (methods) research team (clinical) e-mail discussion mailbase for spm compiler scans published literature web search conference proceedings	papers (from internet) internet database protocol supervisor discussion with colleagues lab meeting

Creating information: some points about data

- ▣ increasing interest from funders, and *some* researchers
- ▣ most researchers spend much of their time searching for, gathering, organising, and analysing data
- ▣ but producing – and sharing - data is not the primary objective
 - ▣ general assumption that data do not have intrinsic meaning until analysed, interpreted, described.....
- ▣ data curation/stewardship/management important to researchers only (at best) intermittently

Different roles and activities: types of information being created

Postdoc	PhD student	PhD student	Research technician
project progress design of experiment SOPs amendment specific techniques demonstration of techniques experimental cell culture techniques repeats of previous experiments analysis of data results visualised under microscope photographs plasmids information for sequencing sequencing alignment of sequencing results with original sequence orders for supplies feedback of some products discuss institute's it system for it committee meeting plan student's project. poster	lab book notes gel image files cell data by microscopy vector nti maps experimental design presentation powerpoint presentations poster notes on other people's work that may be relevant inc. cloning, gene signalling & imaging	lab book entries images of gels from experiments quantitative data from nanodrop equipment emails notes on lab group meeting notes on progress meeting with supervisor notes for an upcoming talk slides for an upcoming talk	lab book writing up daily experiment methods order form

Sharing and disseminating information: ownership, protection and trust

- ▣ data: responsibility, protectiveness and desire for control
 - ▣ lack of rewards for data sharing
 - ▣ concerns about inappropriate use
 - ▣ preference for co-operative arrangements and direct contact with potential users
 - ▣ decisions on when and how to share
 - ▣ commercial, ethical, legal issues

- ▣ belief that only researchers themselves can have the knowledge necessary to take care of their data
 - ▣ intricacies of experimental design and processes
 - ▣ data management plans required by funders, but not much sign of adoption
 - ▣ role of publishers?

- ▣ trust in other researchers' data?
 - ▣ *"I don't know if they have done it to the same standards I would have done it"*

Where, when and how to publish?

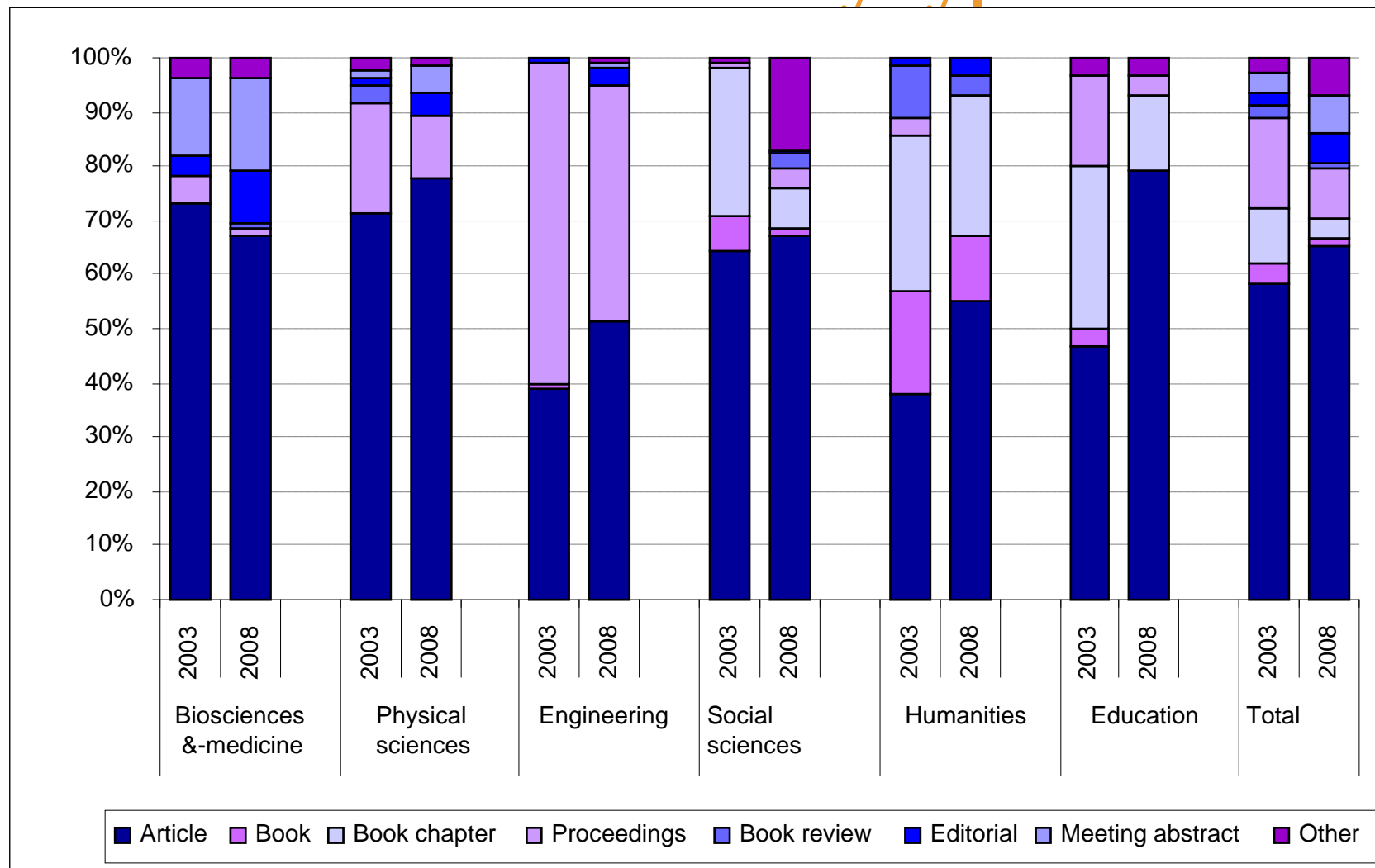
- **key motivations in making choices**
 - register claim
 - maximise dissemination
 - peer recognition (and the rewards that flow from that)
 - some tensions between effective dissemination and prestige
- **publications as measures of performance**
 - growing dominance of journal articles
 - but mixed messages from funders and institutions
- **increasing collaboration** → **more co-authorship**
 - implications for measures of productivity and impact
- **disciplinary differences**
 - monographs in the humanities
 - conference proceedings in engineering and computer science

Productivity?

Average number of publications of UK researchers whose work was submitted to the Research Assessment Exercise

	2003 Mean	2008 Mean	Standard error
bio-medicine	2.32	4.52	.597
physical sciences	4.29	4.51	.609
engineering	3.22	3.31	.474
social studies	2.00	2.69	.266
humanities	1.93	1.70	.124
total	2.50	3.19	.206

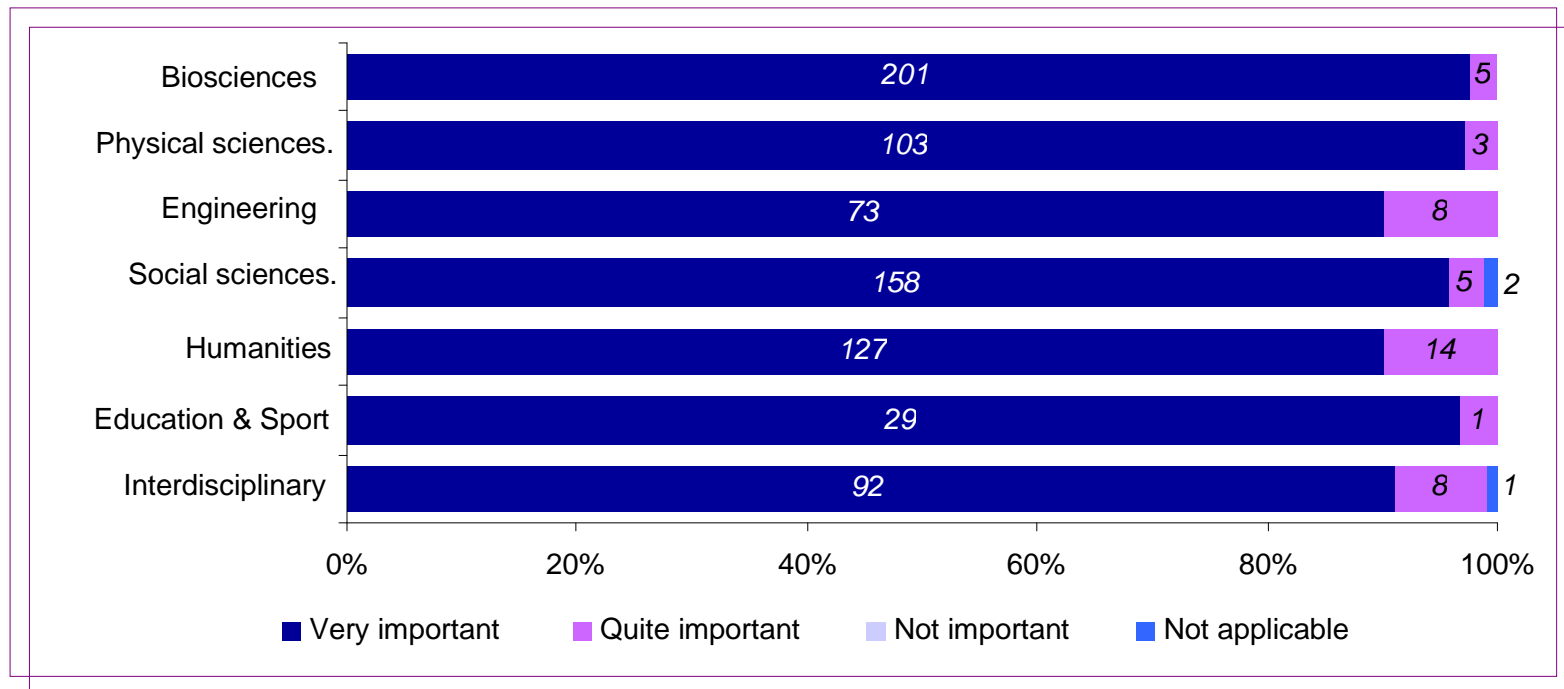
Publications by type



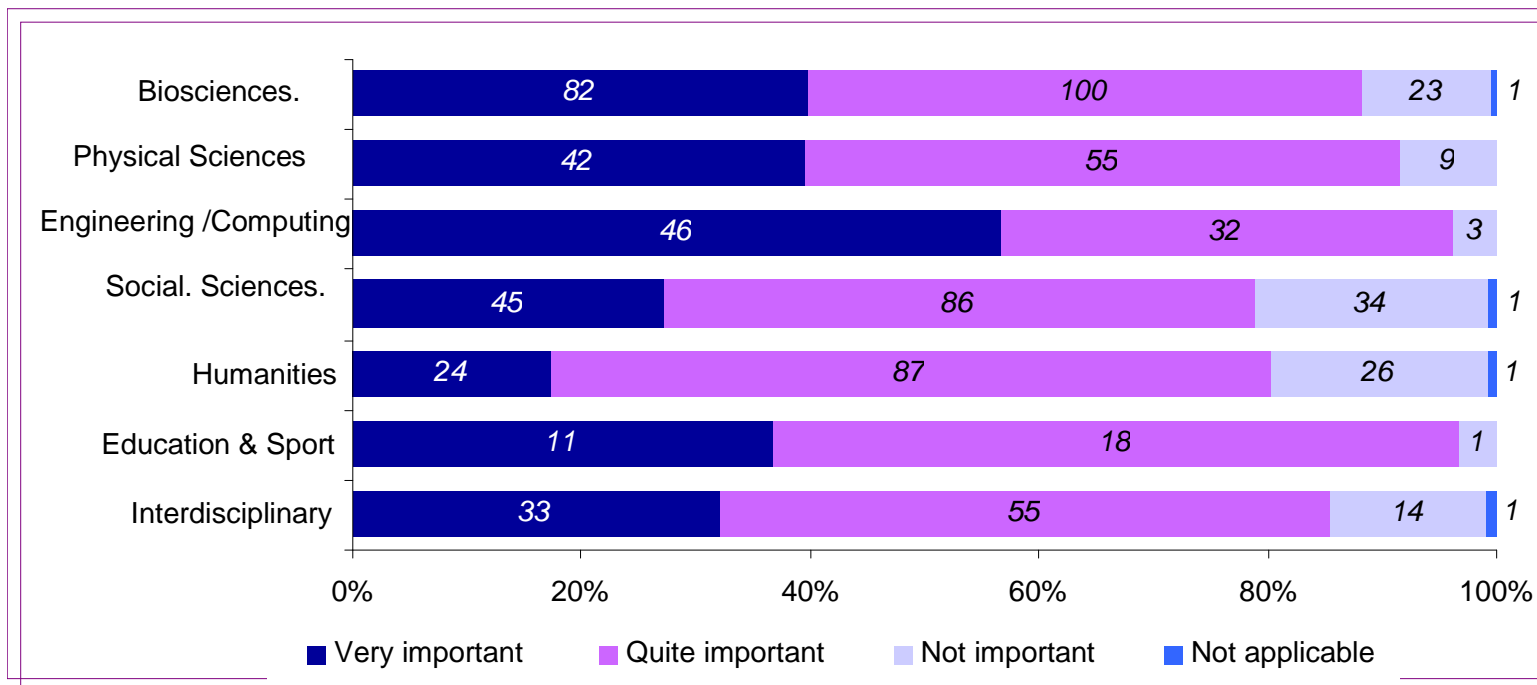
Importance of types of output

	Very important (%)	Quite important (%)	Not important (%)	Not applicable (%)
Scholarly journals	94	6	0.1	0.5
Conference presentations/posters	34	52	13	0.5
Monographs	34	25	32	9
Book chapters	23	60	16	1
Professional journals	19	30	36	14
Works in OA Repository	10	28	41	20
Reports	9	35	44	13
Datasets	8	20	39	33
Working papers	5	27	51	18
Creative works (inc exhibitions & performances)	3	8	40	50
Internet blog/forum	2	10	70	18
Other	7	5	19	70

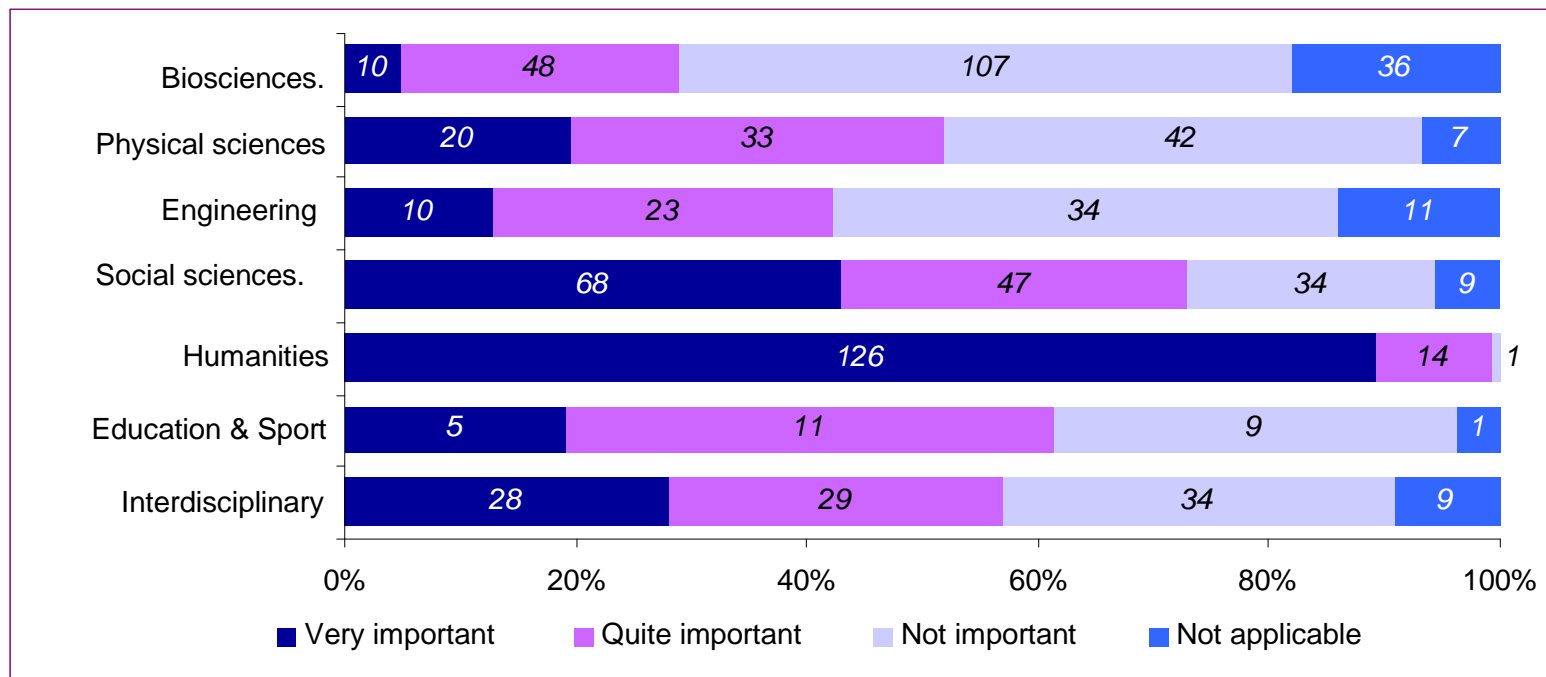
Importance of scholarly journals



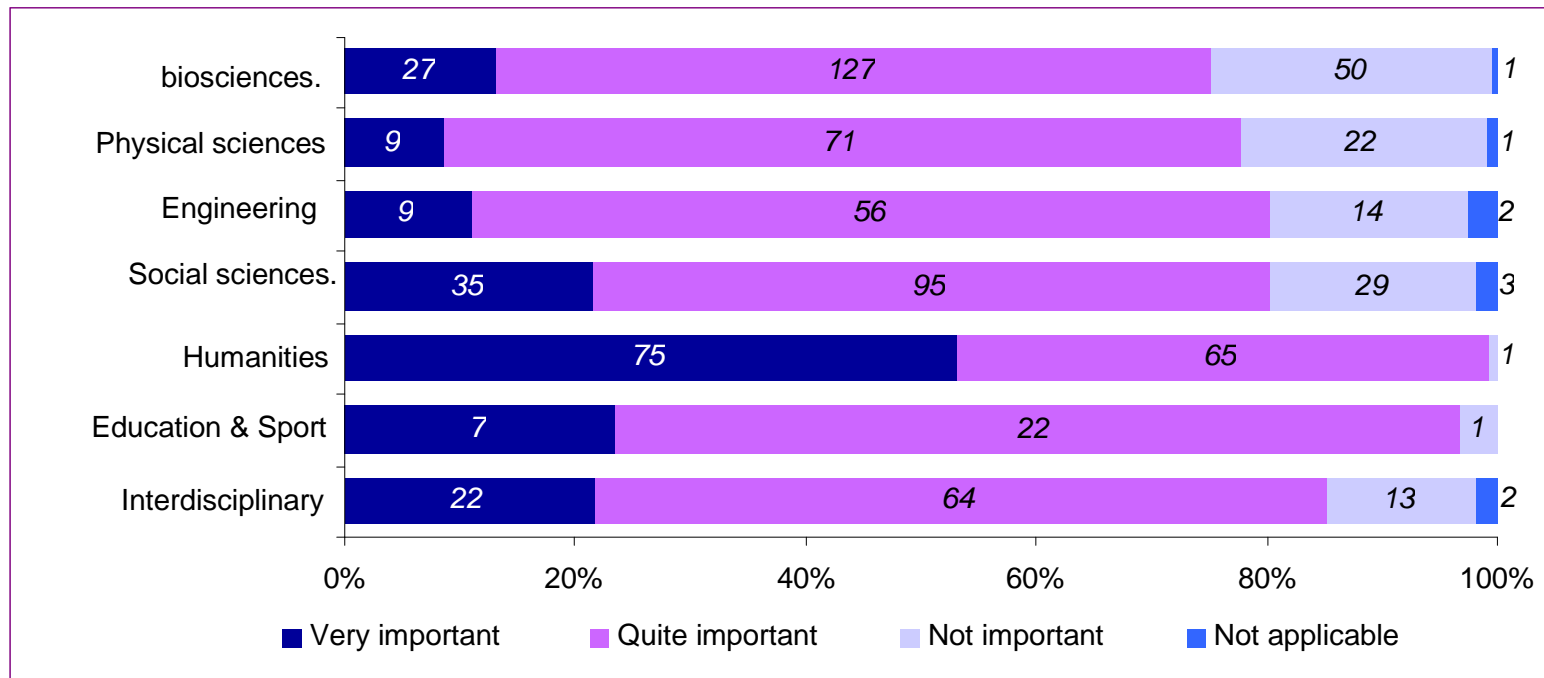
Importance of conference proceedings



Importance of monographs



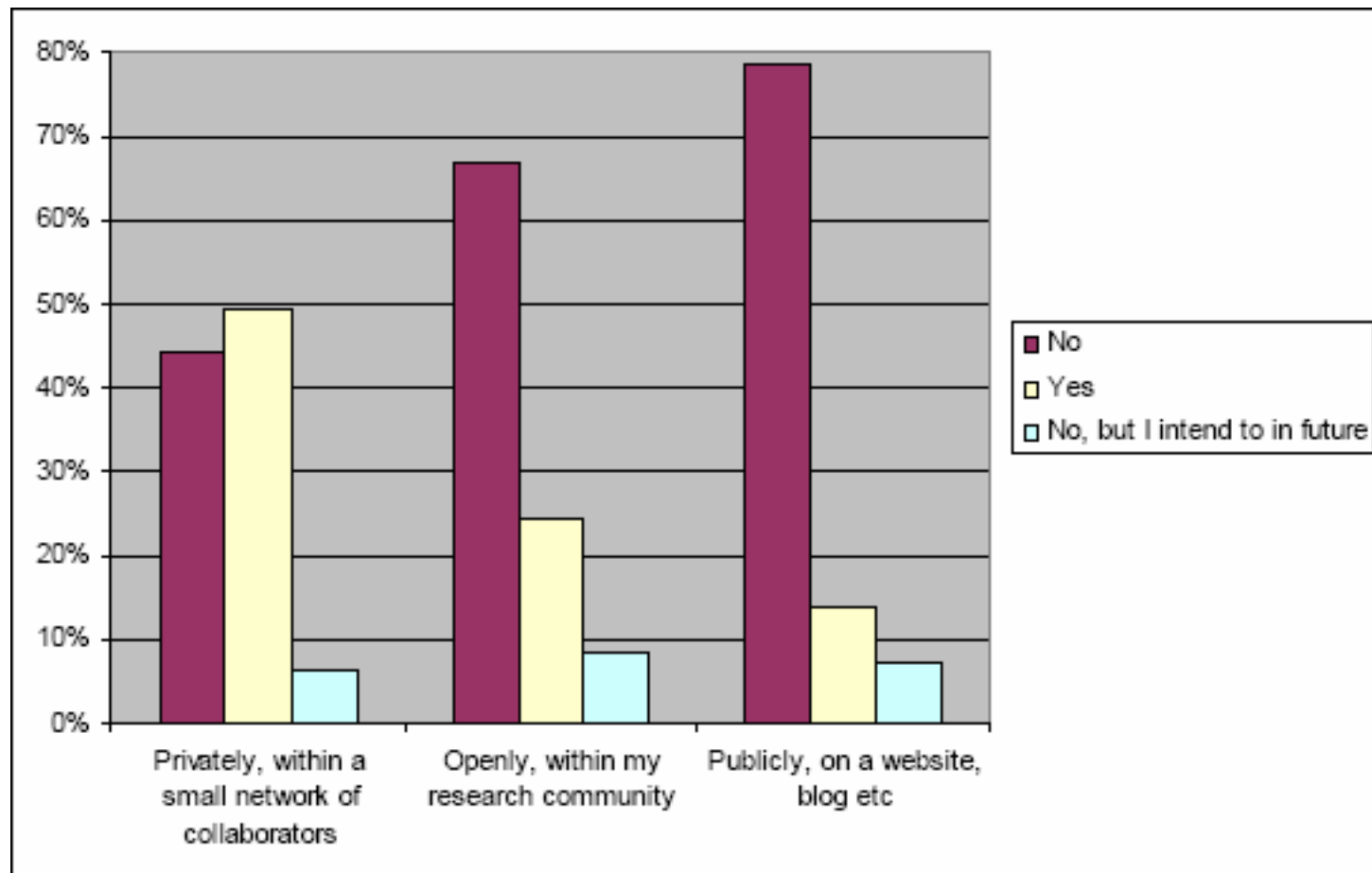
Importance of book chapters



Changes in scholarly communications?

The likelihood of the following changes in scholarly communications within your field over the next 5 years					
	Professor	Reader	Senior Lecturer	Lecturer	Research Fellow
Existing peer review processes will become increasingly unsustainable					
Likely	31%	34%	39%	30%	38%
Unlikely	63%	51%	50%	52%	56%
No opinion	6%	14%	11%	18%	5%
Formal peer review will be increasingly complemented by reader-based ratings, annotations, downloads or citations					
Likely	44%	37%	45%	41%	36%
Unlikely	42%	54%	38%	41%	38%
No opinion	15%	9%	18%	18%	26%
New types of online publication, using new kinds of media formats and content, will grow in importance					
Likely	72%	69%	76%	68%	82%
Unlikely	18%	20%	7%	18%	13%
No opinion	11%	11%	16%	14%	5%
Open access online publication supported by an 'author-pays' funding model will predominate					
Likely	34%	20%	21%	23%	21%
Unlikely	47%	49%	52%	50%	51%
No opinion	19%	31%	27%	27%	28%

Work in progress?



Work in progress?

Do you publish your WORK IN PROGRESS?					
	Professor	Reader	Senior Lecturer	Lecturer	Research Fellow
Privately, within a small network of collaborators					
No	51%	36%	44%	53%	38%
Yes	48%	61%	52%	42%	57%
No, but I intend to in future	1%	3%	4%	5%	5%
Openly, within my research community					
No	76%	79%	79%	86%	76%
Yes	24%	18%	15%	12%	22%
No, but I intend to in future	0%	3%	6%	2%	3%
Publicly, on a website, blog etc					
No	83%	81%	87%	86%	84%
Yes	16%	13%	7%	12%	11%
No, but I intend to in future	1%	6%	6%	2%	5%

Web 2.0?



How often do you do the following in the course of your research activities?					
	Age				
	25-34	35-44	45-54	55-64	over 65
Write a blog					
Never	79%	80%	85%	91%	100%
Occasionally	6%	12%	10%	6%	0%
Frequently (At least once a week)	4%	6%	2%	0%	0%
I do this outside of work	11%	2%	3%	3%	0%
Comment on other people's blogs					
Never	69%	68%	81%	82%	93%
Occasionally	17%	22%	16%	15%	7%
Frequently (At least once a week)	0%	2%	0%	0%	0%
I do this outside of work	15%	8%	3%	3%	0%
Contribute to a private wiki					
Never	80%	75%	78%	85%	86%
Occasionally	18%	17%	17%	14%	7%
Frequently (At least once a week)	2%	8%	4%	1%	7%
I do this outside of work	0%	0%	1%	0%	0%
Contribute to a public wiki (e.g., Wikipedia)					
Never	69%	74%	75%	80%	80%
Occasionally	22%	21%	23%	18%	13%
Frequently (At least once a week)	0%	1%	1%	0%	0%
I do this outside of work	10%	4%	2%	3%	7%
Add comments to online journal articles or more general media publications					
Never	81%	76%	80%	73%	93%
Occasionally	17%	21%	14%	27%	7%
Frequently (At least once a week)	0%	1%	2%	0%	0%
I do this outside of work	2%	2%	4%	0%	0%
Post slides, texts, images, code, algorithms, videos or other media on an open content sharing site					
Never	65%	56%	52%	52%	93%
Occasionally	19%	30%	40%	30%	7%
Frequently (At least once a week)	8%	10%	5%	11%	0%
I do this outside of work	8%	4%	3%	6%	0%

Some tentative conclusions

- researchers vary
 - by discipline
 - by role
- there's still a skills gap
- discovery and access still present challenges
- attitudes towards research data are not always what funders, employers (and publishers?) think they are or should be
- researchers' views of the importance of different types of output do not always correlate with what and how they publish
- Web 2.0 and related developments are small scale as yet, but have the potential to take off

A final thought

- in difficult economic circumstances, researchers will fight harder for funds to sustain their research than for funds to support the information services on which they depend.....



Questions?

Thanks

Michael Jubb
www.rin.ac.uk