Open Access – why and how

Douglas B. Kell

Chief Executive
Biotechnology and Biological Sciences Research Council
(and RC 'Champion' for OA)

Oxford Future of Science Conference
Rigour and Openness in 21st Century Science
April 11th, 2013

Overview of talk

- Open Access – why do we want it?
- Openness and rigour and new opportunities
- Green and Gold OA
- RCUK implementation
- Some current discussion foci
- CC-BY licensing
- Value to be added by semantic enrichment, text mining and related tools

Some reasons for desiring Open Access

- Work paid for from the public purse should be available to the public
- History of market failure of subscription model; authors provide content (and often review) free but cannot then access it
- For researchers, accessible (OA) research is more highly cited
- ‘Public-ation’ is hardly public when behind a paywall
- With 2 peer-reviewed papers per minute being published in BioMedicine (PubMed) alone (~5 overall) only computers can ‘read’ them all
- A free license (such as Creative Commons CC BY) that allows full reuse allows anyone to add value, using techniques such as text mining, semantic mark-up, etc.

‘Openness and rigour’

- OA is a natural part of the normal process of experimental science by which publications must be supported by sufficient information (and access to the underlying data and materials) to permit reproduction of the experiments that led to the findings (and interpretations) claimed.
- So the new digital age does not especially change this; it merely makes it easier and faster and allows better collaboration (which helps with statistical power)

Science’s powerful capacity for self-correction comes from this openness to scrutiny and challenge.

Science as an open enterprise
Royal Society, June 2012.

Ioannides, PLoS Medicine 2, e124 (2005)

Open access, freely available online

Essay

Why Most Published Research Findings Are False

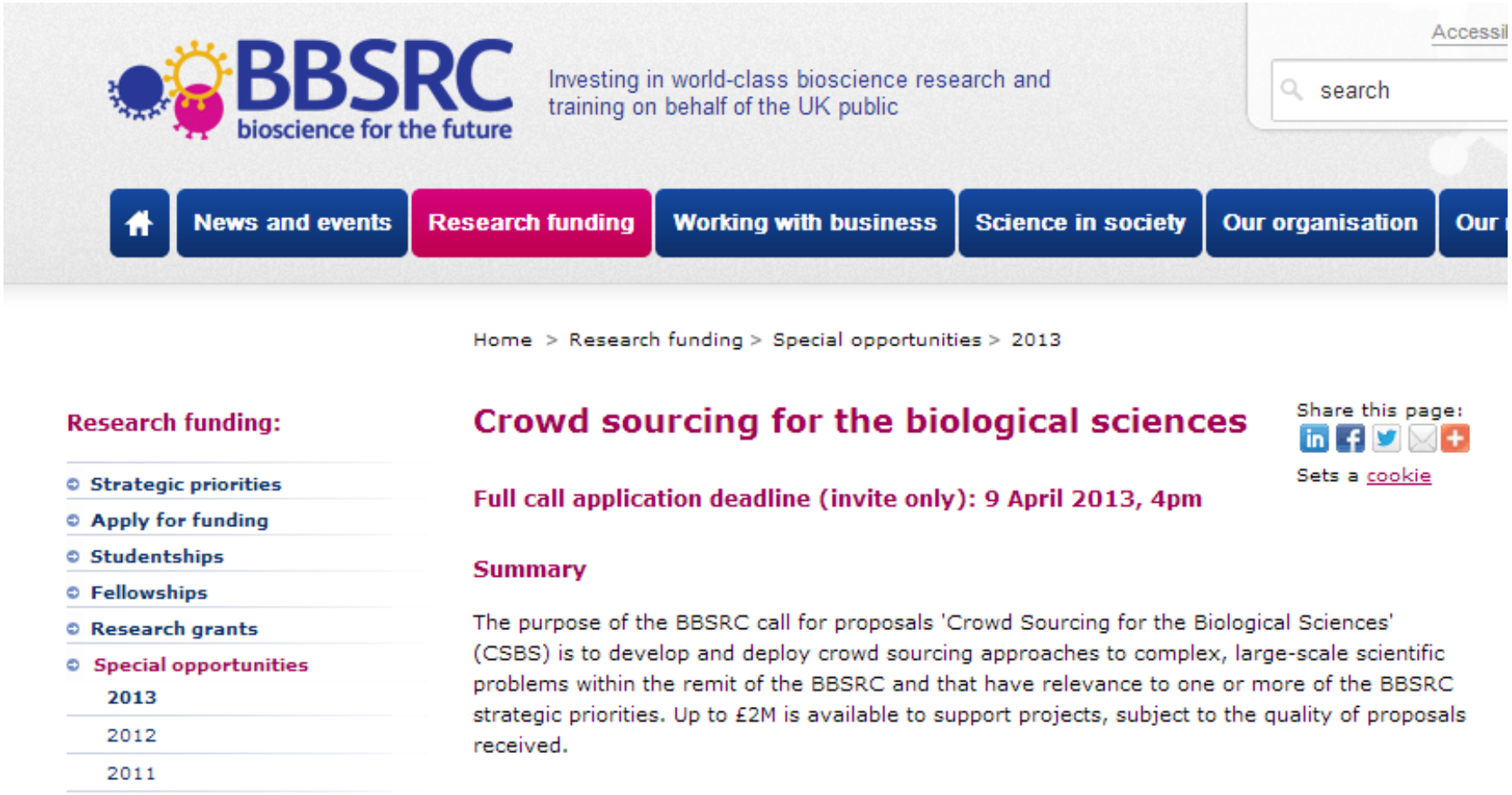
John P. A. Ioannidis

Metabolomics, Vol. 2, No. 4, December 2006 (© 2007)
DOI: 10.1007/s11306-006-0037-z

Statistical strategies for avoiding false discoveries in metabolomics and related experiments

David I. Broadhurst,[a,b,*] and Douglas B. Kell[a,b,*]

BBSRC example of crowd sourcing



International context and examples of OA

- Many countries already implementing OA policies (e.g. Austria since 2004, DFG and ERC since 2006)
- For EC will be part of Horizon 2020
- Science Europe supporting development of coordinated policies
- US NIH enforcing mandate (and provide $100M/y for publication charges; NSF providing $25M/y)
- Other US agencies being mandated to develop OA by OSTP. FASTR proposals would give 6 months embargo
- Global RC has OA as main agenda item in May 2013
- REF consultation giving clear steer to full OA at time of submission/publication of articles

Finch Context: Differing Interests

Different parties have differing interests

- *Universities* …. maximise research performance, control costs
- *Researchers* ……. publish in the best journals
- *Funders* …..maximum impact, control costs
- *Libraries* ……maximise services to readers, control costs
- *Publishers*…..revenues to secure profitability and high quality services/products

Slide courtesy of Prof Dame Janet Finch

‘Green’ and ‘Gold’ OA – RCUK Policy

- Gold preferred by both Finch and RCUK (and Wellcome), involving an ‘article processing charge’ (present average ca £1700), with CC BY licensing allowing full attributed re-use
- If no Gold option offered then Green deposit of final ms after an embargo period, of up to 6 months (biomedicine mandated and STEM) to (initially) 12 months (Arts & Humanities). If Gold offered but funding unavailable then we accept 12/24 months. Anything above 24 months delayed access very much seen as outwith any spirit of OA.
- Assume an initial compliance of 45%
- Most journals of interest are actually compliant now
- Gold OA very widespread, e.g. PLoS One is largest journal, has very effective business model, regards Green as complementary (in assisting dissemination)

“Decision tree”

Recent history of RCUK implementation

- July 2012: Revised RCUK policy on OA launched, alongside HMG response to Finch Report
- September 2012: £10M from BIS to 30 Institutions
- November 2012: RCUK announcement on block grants, £17M in 1st year (from April 2013)
- November 2012 onwards. Many consultations of RCUK with HEIs, publishers, Learned Societies, etc.
- January 2013. HoL Select Committee enquiry on OA
- Early March 2013: Revised RCUK guidance on policy and consultation
- **April 8th, 2013** RCUK Policy and Guidance published at http://www.rcuk.ac.uk/documents/documents/RCUKOpenAccessPolicy.pdf or Google RCUK Open Access Policy
- Q4 of 2014: full evidence-based review of RCUK implementation

Transition to Open Access

- Working with the community to change the way the outputs from Research Council funded work are made available.
- Five year transition to 100% OA – flexibility in implementation.

Journey – not an event

RESEARCH COUNCILS UK

Funding

Arts & Humanities Research Council

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

bbsrc biotechnology and biological sciences research council

£11.2B

MRC Medical Research Council

EPSRC Engineering and Physical Sciences Research Council

NATURAL ENVIRONMENT RESEARCH COUNCIL

Science & Technology Facilities Council

RESEARCH COUNCILS UK

Funding

- Research Councils providing block grants to institutions to support payment of APCs.
- Institutions must establish Publication Funds and the processes and procedures for payment of APCs.
- Flexibility on spend & ‘light touch’ guidance.

Use the money to best deliver the RCUK Policy

RESEARCH COUNCILS UK

Size of the APC fund

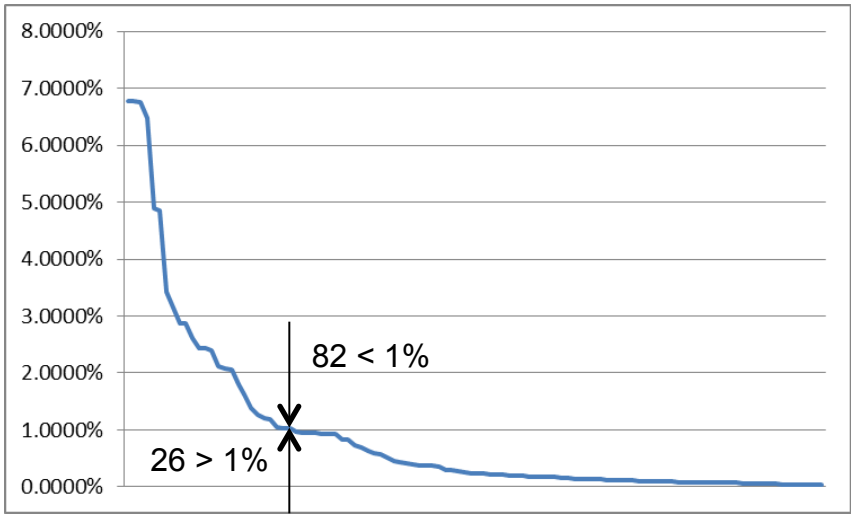
- How many publications?
  - Est. 26k per year, 90% HEI, 10% RC institutes.
- Average APC?
  - Finch £1727 + VAT, paid at 80% fEC = £1658;
- Five-year transition period.

HEI publications	Year-1	Year-2	Year-3	Year-4	Year-5
Est. % Gold	45%	53%	60%	67%	75%
APC fund	£17M	£20M	tbc	tbc	tbc

RESEARCH COUNCILS UK

Distribution of APC fund

- Based on % share of direct labour funding received over past 3 years (£1.5B)
  - *DI Staff and DA Investigators*
- Russell Group & 1994 Group – 37 HEIs, 82%
- Cut off below £10k in year-5 (>99% RC spend)



RESEARCH COUNCILS UK

Supporting the Transition

- Working with Sherpa-Romeo, JISC and Wellcome Trust to develop journal compliance web site.
- Working with the RIN on ‘best practice’ project to develop protocols between HEIs.
- Plans to facilitate workshop for Learned Societies to share ‘best practice’ in OA publishing.
- Revised guidance and information on transition flexibility (April 8th).
- Q4 2014 – evidence based review of policy and its implementation.

RESEARCH COUNCILS UK

CC BY licensing

- Mandatory version for Gold OA when APCs paid
- Allows full re-use, including commercial, with attribution
- Hence is not ‘plagiarism’
- Simplest method for allowing re-use (Hargreaves ++)
- The bounds of ‘non-commercial’ are rather unclear – are Universities ‘commercial’?
- Does not affect third-party rights e.g. copyrighted images (or proprietary software) used in original article with permission remain copyrighted or proprietary
- In common use now

Transparency requirement

- Acknowledgement of funding.
- Statement on access to the underlying research materials.
- Helps support the transparency, integrity and robustness of the research process.

RESEARCH COUNCILS UK

Why full papers, and not only abstracts?

- A survey[1] of 29 biomedical papers showed that authors reported in the abstract fewer than 8% of the scientific claims that actually appeared in the body of the paper
- Of course most abstracts are also deficient in numerical details of the data

[1]Blake C: Beyond genes, proteins, and abstracts: Identifying scientific claims from full-text biomedical articles. J Biomed Inform 2010; 43:173-189.

Text mining – 3 main stages

- **Information retrieval** – finding material that is relevant to the question of interest – needs OA
- **Information extraction** – fact retrieval – adds value
- **Data mining** – with 'deep' parsing and semantic annotation this allows true text mining – creates knowledge

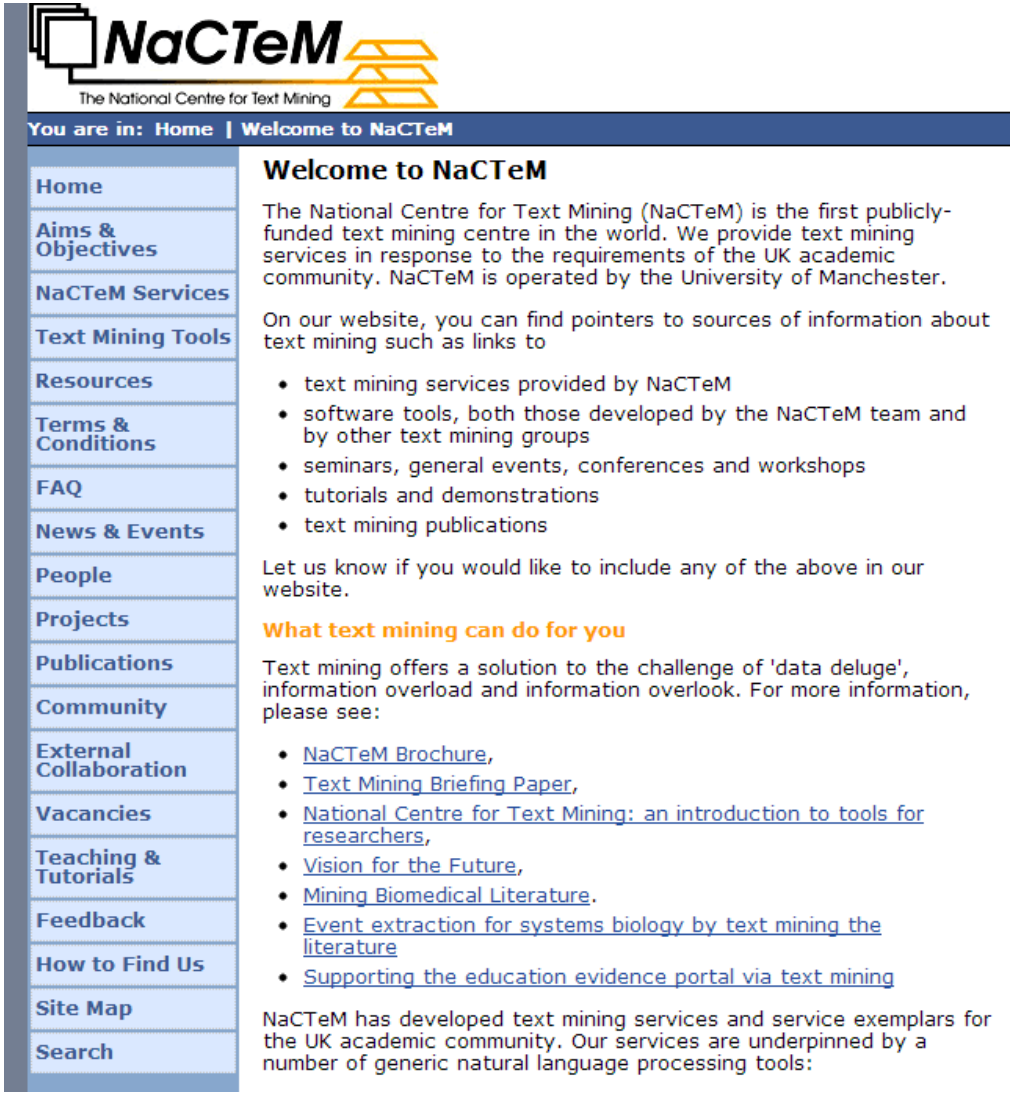
Text mining and its potential applications in systems biology

Sophia Ananiadou[1,2], Douglas B. Kell[3,4] and Jun-ichi Tsujii[1,2,5]

Trends Biotechnol 24, 571-579 (2006)

National Centre for Text Mining (NaCTeM)

- http://nactem.ac.uk/
- Various tools available
- Mainly abstracts
- Full text where OA



Current tools

NaCTeM has developed text mining services and service exemplars for the UK academic community. Our services are underpinned by a number of generic natural language processing tools:

- TerMine is a Term Management System which identifies key phrases in text.
- AcroMine is an acronym dictionary which can be used to find distinct expanded forms of acronyms from MEDLINE.
- Kleio is an advanced information retrieval system providing knowledge enriched searching for biomedicine.
- FACTA+ is a MEDLINE search engine for finding associations between biomedical concepts.
- IRS facilitates advanced searching of documents by making use of added value features extracted from full texts using NaCTeM text mining tools.
- MEDIE uses semantic search to retrieve biomedical correlations from MEDLINE.
- Info-PubMed uses a gene/protein dictionary and deep parsing to understand protein interactions **[Firefox Required]**.

BMC Medical Genomics

BioMed Central

2,469 refs

Review

Open Access

Iron behaving badly: inappropriate iron chelation as a major contributor to the aetiology of vascular and other progressive inflammatory and degenerative diseases

Douglas B Kell*

http://www.biomedcentral.com/1755-8794/2/2/

http://dbkgroup.org/publications/

Arch Toxicol (2010) 84:825–889
DOI 10.1007/s00204-010-0577-x

Open Access

REVIEW ARTICLE

1,716 refs

Towards a unifying, systems biology understanding of large-scale cellular death and destruction caused by poorly liganded iron: Parkinson's, Huntington's, Alzheimer's, prions, bactericides, chemical toxicology and others as examples

Douglas B. Kell

Concluding remarks

- Implementation of Finch recommendations proceeding with momentum and money
- Preference for Gold / CC BY to allow immediate OA and full re-use, but a mixed economy (with Green) accepted
- A journey rather than a fixed point
- Strong international context
- Need to modernise elements of copyright (BIS document in December); significant discussions within EC
- Huge **opportunities** in adding value and novel digital enhancements to OA texts for imaginative publishers and other entrepreneurs; many have begun to realise them

Open Access – why and how

Douglas B. Kell

Chief Executive
Biotechnology and Biological Sciences Research Council
(and RC 'Champion' for OA)

Oxford Future of Science Conference
Rigour and Openness in 21st Century Science
April 11th, 2013