

UK PubMed Central: Interview with Phil Vaughan

Martin Fenner, Gobbledygook

November 4, 2009

PubMed Central was launched in February 2000 by the U.S. National Institutes of Health (NIH) as a free digital archive of journal articles. Just as PubMed, PubMed Central covers research in the life sciences, but not other areas of research, e.g. engineering, physical sciences or astronomy.

Some journal articles are available as full text as soon as they are published, and most journals provide free access to full text articles within a year of publication. Some journals only provide the full text of some articles, including research funded by the NIH under the NIH Public Access Policy. The majority of fulltext articles in PubMed Central are not Open Access, but are protected by copyright. These articles are often made available under a license that allows redistribution and reuse.

All articles are deposited using the NLM-DTD XML format, which is a standard text format suitable for text mining and long-term archiving. Most articles are deposited directly by the journals, so that authors do not have to get involved in the technical aspects of article deposition.

PubMed Central is a centralized archive of full text papers and not simply an interface to search these articles at the websites of the participating journals. Neither is PubMed Central an interface to search the various institutional repositories at universities and institutions. The NIH thinks that this centralized approach makes it easier to develop additional functionality, including the integration with other databases (e.g. the protein or nucleotide databases) hosted at the NIH.

UK PubMed Central was launched in 2007 as the first PubMed Central outside the United States (PubMed Central Canada, the second international PubMed Central, launched this week). In January 2010 UK PubMed Central will launch a number of new services, and I used the opportunity to ask UK PubMed Central Programme Manager Philip Vaughan a few questions.



Figure 1:

1. What is UK PubMed Central?

UK PubMed Central offers free access to 1.6 million full text journal articles in the fields of Life Sciences, Biomedicine and Health. Our content is free to read, print off and download. There is no registration process required: users can simply visit our site and start searching! It has been developed in consultation with the UK biomedical and health research community, and is about to launch some exciting new services in January 2010. Usage of our service has been growing steadily; we average 300,000 downloads a month currently.

2. Why do we need more than one PubMed Central, and how are they connected

UKPMC gives the UK research community access to all the content of PMC, but with added value specifically for the UK. For example, it offers Grant Reporting features, whereby users can search for current grants from all of our 8 Funders, can link the publications they produce to the grants they originated from, and view the impact of their work through access to citations. UK researchers can also deposit their manuscripts directly into UKPMC. We will also be accessing additional content not in PMC; around 475,000 extra articles, reviews, guidelines and theses. Through the involvement of our funders, we are capturing around 90% of recently published journal articles from biomedical research conducted in the UK. We are therefore increasing the visibility of UK research in the field.

Our colleagues at NCBI in the US have been very supportive of our activities: they have been keen for the PMC “project” to expand beyond the US. A Canadian version (PMC Canada) launched this week.

3. Why does a search use PubMed Central and not the UK version?

The current UKPMC site searches PMC as it “mirrors” its content. But this about to change; from January UKPMC will be a “stand alone” service with its own up to date archive of content, as well as access to all the content of PubMed and PubMed Central.

4. What is the relationship between UK PubMed Central and institutional repositories?

UKPMC is a subject repository and as such receives content from researchers at Higher Education institutions across the UK. Deposition in UKPMC is mandatory if a researcher has a grant from any of our 8 Funders: Wellcome Trust, Cancer Research UK, Medical Research Council, British Heart Foundation, Arthritis Research Campaign, Biotechnology and Biosciences Research Council, National Institute for Health Research and the Chief Scientists Office. Consequently we do not need to harvest content from other repositories. However, institutional repositories are welcome to harvest our content; we have an OAI-PMH interface to enable this. On the other hand, there are no plans to develop of institutional repositories from within UK PubMed Central.

5. What is the material that can't be put into UK PubMed Central?

If the material is not peer-reviewed research published in a recognized journal by a known publisher, it would not be deposited into UK PubMed Central. One aspect of our development programme has been to identify what further content we could provide links to, for example published research theses and NHS clinical guidelines. But, we have rigid quality control procedures in place even to provide links to material. That said, we do remain open minded, in terms of what material we may consider linking to in the future. And who knows what might arise in the future. For example, there does seem to be an appetite to include images.

6. Can I submit my accepted manuscript if it was not funded by one of the UK PubMed Central funders?

Unfortunately not at this time. We are hopeful that other UK research councils and funders of life sciences research will come on board in due course. The more funders that come on board, the stronger our service will become and the stronger our message on the importance of OA.

7. What is the UK PubMed Central OAI service?

The UK PubMed Central OAI service, (UKPMC-OAI) provides access to metadata of all items in the UKPMC archive, as well as to the full text of a subset of these items.

8. What are your responsibilities at UK PubMed Central?

I am the Programme Manager. I am responsible for the current service which has been in existence since January 2007. I also co-ordinate all our current development activities, which are undertaken by ourselves at the British Library, by the University of Manchester (MIMAS), the National Centre for Text Mining (NacTEM – based at the University of Manchester) and the European Bioinformatics Institute (EBI) based at Hinxton in Cambridgeshire.

9. What did you do before starting to work on UK PubMed Central?

I originally trained as a librarian. Since then I have worked mainly in health and medical information work, both in Higher Education and in the National Health Service. My most recent post was for JISC (Joint Information Systems Committee) as a Programme Manager. I was responsible for a portfolio of development projects in the digital library sphere. I joined UKPMC at the British Library in May 2008.

10. Do you want to talk about future plans for UK PubMed Central?

Yes please! We are soon to launch some exciting new initiatives in January 2010. Our website has been completely redesigned with a new more intuitive interface,

and we will be expanding our content to include all content from PubMed (c. 18m references), patents, theses and NHS Clinical Guidelines. Consequently our users will be able to access a vast collection of relevant content through one portal.

Our search engine will be utilising innovative new techniques such as text mining to retrieve more contextually relevant information, link to other relevant databases and provide a richer search experience for the user. We are also expanding our Grant Reporting functions to allow Grantees and Funders to assess the impact and value of their funded research, and increase their visibility.

Our Funders are keen to expand the programme, possibly towards a future European PubMed Central. We are hopeful that as part of this process some European Funders may be joining the programme shortly. Our Funders hope to make an announcement regarding this shortly.

We are holding a Showcase event on January 12th at the British Library in London to highlight these new developments and the future potential of the service.

11. Could you provide contact information for people that have further questions about PubMed Central?

By all means, people can email us at ukpmc@bl.uk and my email is Philip.Vaughan@bl.uk.