The inclusion of an online journal in PubMed central – a difficult path

Victor Grech

To cite this article: Victor Grech (2016): The inclusion of an online journal in PubMed central – a difficult path, Journal of Visual Communication in Medicine, DOI: 10.1080/17453054.2016.1182472

To link to this article: http://dx.doi.org/10.1080/17453054.2016.1182472

Published online: 31 May 2016.

Submit your article to this journal

View related articles

View Crossmark data
The inclusion of an online journal in PubMed central – a difficult path

Victor Grech

Academic Department of Paediatrics, Medical School, Mater Dei Hospital, Malta

ABSTRACT

The indexing of a journal in a prominent database (such as PubMed) is an important imprimatur. Journals accepted for inclusion in PubMed Central (PMC) are automatically indexed in PubMed but must provide the entire contents of their publications as XML-tagged (Extensible Markup Language) data files compliant with PubMed’s document type definition (DTD). This paper describes the various attempts that the journal Images in Paediatric Cardiology made in its efforts to convert the journal contents (including all of the extant backlog) to PMC-compliant XML for archiving and indexing in PubMed after the journal was accepted for inclusion by the database.

ARTICLE HISTORY

Received 29 September 2015
Revised 17 December 2015
Accepted 2 March 2016

KEYWORDS

Education; multimedia; Printed media

Introduction

Having a journal indexed in PubMed or PubMed Central is an onerous task that is not only fraught with difficulty but also requires multiple practical steps. This paper will illustrate the route by which the journal Images in Paediatric Cardiology was accepted for indexing in PubMed Central, and hence in PubMed. It is hoped that this paper will guide editors (and their assistants) should they wish to attempt to index a journal in this prestigious database.

PubMed and PubMed central

The indexing of a journal in a prominent database is an imprimatur that almost automatically labels such a journal as a source of good-quality information and gives it the potential to further its authority in the field. Indexing and abstracting also facilitates the dissemination of information published through such journals to researchers within these journals’ particular fields.

It is considered appropriate for a journal to attempt indexing and abstracting within such a database when it has a solid track record for timely publication and relevant content. There are many such databases, and within the biomedical literature, PubMed is arguably the most coveted database.

PubMed is a service offered by the US National Library of Medicine (NLM) and currently includes over 23 million citations from over 5,500 journals for biomedical articles that date back to the 1950s, abstracted MEDLINE and other life sciences journals. PubMed also includes links to full-text articles and to publishers’ websites and is searchable, thereby serving as a premier online educational and research facility (Grech, 2009).

MEDLINE is NLM’s database of indexed journal citations and abstracts, covering nearly 4500 journals internationally at the time of writing. Fields covered include medicine, nursing, dentistry, veterinary medicine, bioengineering, chemical and preclinical sciences. MEDLINE includes references to articles indexed as far back as 1966, and even older citations are available in an NLM database known as OLDMEDLINE.

Information technology and the Internet have yielded storage and searching technologies that permit users to access abstracts or even full-length articles at the click of a mouse button, such that MEDLINE has become the default online search system for health and allied sciences. MEDLINE also has an advanced search facility with specific Boolean filters, thereby facilitating advanced searches. The service is free, and is maintained by the National Centre for Biotechnology Information (NCBI) at the National Library of Medicine (NLM) at the US National Institute of Health (NIH). Due to the breadth and scope of PubMed, journals (and therefore articles contained therein) that are not indexed in this database are at an instant disadvantage in terms of exposure to readers and researchers.

The NLM utilises the NIH-commissioned Literature Selection Technical Review Committee (LSTRC) to review applications from journals that wish to be indexed in PubMed. The criteria for a journal to be included in the database, thereby gaining repute and cachet, are several and include journals that routinely
expose most of their content to the peer-review process (Albert, 1999). However, there is, in addition, a large body of so-called ‘grey literature’ that is not peer-reviewed, but is also included.

The literature selection technical review committee (LSTRC) routinely requires a formal application from applying journals along with several issues for evaluation and vetting purposes. The main criteria that are sought are quality, originality and importance of the scientific content. The journal must have been published for six months, with at least 20 articles available. The content must all be digitally archivable. About a quarter of submitted journals are accepted for indexing, and journals that are turned down may reapply after two years.

Accepted journals must provide citation and abstract data as XML-tagged (Extensible Markup Language) data compliant with PubMed’s stringent requirements, according to a pre-specified document type definition (DTD).

XML is a standard format for embedding descriptive information that can be automatically included as fields in a database. This forces each and every individual element (including such detail as authors’ initial/s) to dovetail into a logical, predictable and hierarchical structure.

The NLM encourages journals to make all content available gratis by offering a free archiving service in PubMed Central (PMC), thereby also allowing researchers free access to all such content (Delamothe, 1999). This initiative commenced in February 2000 and is managed by NCBI. Free and permanent access to researchers is a core principle of PMC, and the project is committed to providing permanent access to all content, even as technology evolves and old formats threaten to become obsolete. Copyright is retained by the journal publishers and individual authors (Grech, 2001b). Conversely, journals accepted for archiving in PMC are automatically indexed in PubMed. Such journals must submit content in XML compliant with PMC’s DTD.

The advantages of such online publications are numerous and include lower costs for journals, the ability to include multimedia and a more rapid publishing turnaround time (Grech, 2001a).

Images in paediatric cardiology

*Images in Paediatric Cardiology* was conceived by this author in 1999 (Ellul-Micallef, 1999), when the potential of the Internet was becoming apparent (Grech, 2001d). An international editorial board was created and said author began personally soliciting articles. The journal was initially hosted by the Maltese Department of Health in a subdomain and was later moved to a server and hosted by a sponsor (Shadowservices), and a domain name was purchased (www.impaedcard.com).

This author was solely responsible for maintaining the journal’s online presence through hypertext, such that each paper continues to be published not only as a standard HTML web page but also as PDF (Grech, 2002), along with graphic editing (Grech, 2000), scanning (Zammit, 2001) and animation editing (Grech, 2003). The peer review of individual articles is carried out in the customary manner by the editorial board and invited reviewers, and this author is responsible for technical reviews.

The journal is free to all and has been of an open-access nature from well before the term was actually coined (Varmus, 1999), and as is customary in open-access, copyright is retained by the author/s (Grech, 2001c).

Inclusion in PubMed central

*Images Paediatric Cardiology* applied for inclusion in PMC in 2003. Acceptance came in that same year and it was then up to the journal, with its very limited resources, to make the journal available in XML compliant with PMC’s DTD.

Direct funding was unavailable as the journal, up to that date, did not charge authors for the inclusion of their accepted articles in a journal issue. Furthermore, no revenue is generated to date from advertising and, being open access, no revenue is generated from subscription fees.

A medical colleague with strong IT skills (Dr. Joe Pace) attempted to create a parser in 2006 order to convert the journal from its native HTML to XML, but this proved technically unfeasible and unwieldy and the attempt was abandoned.

The (local) University of Malta attempted to help in two ways. Attempts were made to locate European Union (EU) Funds for XML conversion in 2007 (Sir Temi Zammit Foundation), but these attempts did not yield fruit. Similarly, the University’s IT Department also promised to allocate a student to create the requisite software for the conversion process, but these efforts also came to naught.

A local (Alert Communications) IT company and personal friend of the author similarly strove to obtain EU funding in 2010 and promised to sponsor any shortfalls, but these attempts to obtain funding also failed. The editor therefore approached various pharmaceutical and non-pharmaceutical companies and amassed a ‘war-chest’ for the conversion of all journal issues to XML.

An Internet search indicated that eXtyles software (Inera) would be able to convert to XML. This is an integrated suite of tools that allows editorial staff to work in the Microsoft Word environment. It is also advertised as simplifying the time-consuming and
Table 1. A typical XML example. The output is shown in Figure 1.

Table 1

| Tetralogy of Fallot with absent pulmonary valve syndrome; an imaging challenge |

Figure 1. Output of code in Table 1.

The Malta medical journal

Malta has two other peer-reviewed biomedical journals, the Malta Medical Journal (MMJ, http://www.um.edu.mt/umms/mmj ISSN: 1813-3339, EISSN: 2308-4103) and the Journal of the Malta College of Pharmacy Practice (http://www.mcppnet.org/publications.htm). This author has been the Editor of the former journal since 2013. Word documents containing the complete text, as well as separate high-quality files for images, have been archived. For the period prior to this date, going as far back as 1988, the journal has been scanned to PDF with optical character recognition.

The MMJ has applied for inclusion in PMC at http://www.ncbi.nlm.nih.gov/pmc/utils/publisherportal/apps. This is the US National Library of Medicine’s New Journal Application portal for participation in PMC. Applications are open to any English-language life sciences journal that meets NLM’s standards for the archive. Applicant journals must qualify on two levels: scientific content and technical quality. As explained earlier, accepted journals must provide full-text XML and any accompanying media and indexed in PMC, it became possible to charge a small fee in order to maintain funding for conversion of further issues, the purchase of hardware and software, hosting and domain name registration.

The journal is selective and now publishes only interesting images, thereby adhering to the journal name in fact as well as in spirit. It has been a long and serendipitous journey but well worth the effort.
supplementary material files to PMC for archiving. Exeter Premedia Services have in principle agreed to convert the journal to XML compliant with PMC’s DTD and funding has been made available by the University of Malta to support this initiative. The MMJ’s editorial board eagerly awaits PMC’s decision.

Conclusion

The factors required to index a journal can be summarised thus:

1. The availability of funding in order to convert papers (including potentially old content) to XML compliant with PMC’s DTD.
2. A publisher who can take on this conversion at mutually agreed rates of conversion and pricing.
3. The submission of a formal application to PubMed central which will result in the applicant journal being vetted for scientific content and technical quality.
4. The submission of full text sample XML and any accompanying media and supplementary material files to PMC by the publisher for testing by PMC.

It is sincerely hoped that this paper will help potential journal entries into PubMed central.

Disclosure statement

The author reports no conflicts of interest. The author alone is responsible for the content and writing of this article.

References