

DRAFT -- To be published in DMT'07 Proceedings
(see <http://ngmdb.usgs.gov/Info/dmt/>)

Open Access Journal Publication: Implementation, Copyright and Dissemination, Using the Journal of Maps as a Case Study

By Mike J. Smith

School of Earth Sciences and Geography
Kingston University
Penrhyn Road
Kingston upon Thames, Surrey, KT1 2EE
Telephone: +44 (207) 099-2817
Fax: +44 (870) 063-3061
email: michael.smith@kingston.ac.uk

INTRODUCTION

Open access (OA) journals are rapidly becoming an important channel for publishing academic articles (Rightscom, 2005) and, although they represent a small proportion of the total number of journals published annually, it is significant that organisations such as British Medical Journals (BMJ) operate in this manner. This article explores the broad implementation of OA journals, issues pertaining to copyright and the distribution of (geospatial) research data.

Much of the material elicited in the preparation of this article was derived from experiences as editor of the Journal of Maps (JoM; <http://www.journalofmaps.com>) and therefore the discussion should be understood within this context. The Journal of Maps was established against the backdrop of a perceived decline in the publication of research based maps (Smith, 2005a). With the movement of print published journals towards a standardised A4 copy format, large maps are not easily publishable. The inclusion of “inserts” (folded or stitched) into journals appears to have declined over the last century (Smith, 2005a) and, with the high cost of colour printing, there is an apparent decline in research map publication. Maps are also rarely seen as a research goal in their own right, with the focus of journal publication often upon the communication of research results. JoM was therefore founded as a charity with the specific remit of publishing research maps.

OPEN ACCESS

Open access can be defined as material that is free *at the point of consumption*. Although simple in concept in that you “give” content away, there are a variety of implementations currently used by journal publishers. There are also hybrid publication models, being part OA and part “paid-for”. Rightscom (2005) summarised the main publishing models currently used for academic publications in the UK, highlighting the growing importance of OA, as well as listing the main methods that are currently implemented.

Research Councils UK (2005) state that over 60% of university research in the UK is funded directly by the government. The dissemination of research results is traditionally performed through academic journals, after peer review has taken place. Indeed, editorial boards

that make up the academic component of journals are normally non-stipendiary positions, indirectly funded by the institution at which they are employed. This work is deemed to be “scholarly activity” and part of the duties of an academic. The journal publishers, however, are interested in making a “reasonable profit” from the publication of a journal title, whilst the “consumers” of journals are typically research institutions (universities). The position that the funding body (in this instance the government) finds itself in, is paying to access the results (through a subscription based publishing model) of research it has already funded. This position is considered untenable by the UK government (HMSO, 2004) and it recommends greater access to research findings, possibly through the lodgment of results in institutional repositories. This position was subsequently backed by the principal research funding bodies, represented by Research Councils UK (RCUK, 2005), who now require the deposition of the results of all funded projects in research council specific repositories (e.g. Natural Environment Research Council; <http://www.neodc.rl.ac.uk/>). Such moves have been partially mirrored in the USA where the National Institutes of Health (NIH) have announced a similar move, requiring the deposition of results from funded work in PubMedCentral (<http://www.pubmedcentral.nih.gov>).

The government initiatives outlined above will clearly be driving the agenda for the establishment of OA journals over the next few years. Other effects of a subscription-based journal publishing model are the restriction of access to current research for developing nations, as well as the increased financial burden placed upon libraries. However it is important to note that journals will never be free to publish (Morris and Powell, 2005), as there are always costs associated with review and publication. Rowland (2002) estimates a cost of \$200-400 per article, so the main concern is whether journals can be *appropriately* funded.

OA does therefore **not** mean a “no cost” publishing model. To re-iterate, it is free at the point of consumption. Users (or consumers) of the material do not have to pay to access the material. Funding sources for OA journals are therefore required from elsewhere. Two broad categories can be identified:

1. Author Pays

This is the most common funding model and has been adopted by large volume OA publishers such as BioMed Central (BioMed Central, 2004) and the Public Library of Science (<http://www.plos.org>). The financial result of this model is that it shifts the cost of article publication from the consumer to the grant body, which may be a research council or the host institution, however it produces barriers to those with little funding. This can be mitigated against with free submission for low income groups.

2. Those that can afford

This has been implemented by the BMJ (Delamothe and Smith, 2003), who previously operated an OA publication methodology, cross-funded by other activities of the organisation. This has not proved sustainable and it now charges “some users some of the time.” Articles are freely distributed immediately upon publication (1 week), followed by a period (1 year) where charges are levied against users from wealthy nations. Developing nations are not charged. The importance of this methodology is that it charges those nations that can afford to subscribe, whilst still freely distributing material elsewhere (defined by the World Bank’s list of 120 low and lower middle income countries).

COPYRIGHT

This section discusses copyright, within the context of JoM, and is based around two issues. The first relates to the copyright that JoM claims based upon the material it publishes. The second relates to third party data included within maps that are published by JoM.

Publishing Copyright

It is common when submitting a research article for publication for the author to retain full copyright up to the point where it is accepted for publication. Many journals, after accepting the article, will require the submission of a copyright transfer form that assigns full, irrevocable, copyright for the material to the publisher. Some authors are unhappy at the requirement to transfer the copyright of their own work to a third party and, for government employees, a separate agreement is often used.

JoM allows the author to retain full copyright, whilst granting JoM an irrevocable license to publish the material. In essence the author and publisher share the copyright. JoM is also open access, with a requirement to make material freely distributable. It has therefore adopted the Creative Commons licensing model (<http://www.creativecommons.org>) that allows detailed specification in the use of published material, whilst still retaining copyright (Creative Commons, 2005). Specifically, the license allows the freedom to copy, distribute, and display all published material for non-commercial purposes, whilst requiring full attribution in its use and non-alteration.

Third Party Copyright

The incorporation of third party material (e.g. photo, diagram, table) in a journal article requires agreement from the owner of the copyright. Within the context of JoM, this principally means the incorporation of third party data within published maps. The copyright restrictions are dependent upon the data supplier and will vary from organisation to organisation. For example, the USA federal government cannot claim copyright (Copyright Office, 2007) for data that has been acquired through the use of public funding and has a mandate to make it available at the cost of distribution. This includes data sets such as the Shuttle Radar Topography Mission digital elevation models and Landsat ETM+ satellite imagery. As these products are public domain, they can be incorporated into other materials.

Within the UK the national mapping agency, the Ordnance Survey (OS), is the primary supplier of geospatial data. Detailed licensing restrictions for universities cover the use and reproduction of data in print and electronically for local and international distribution (EDINA, 2007). In particular, the license is flexible with respect to use of data in posters, presentations, teaching materials, and internal use. However, greater restrictions are imposed upon electronic, publicly accessible, publications; these are based upon a combination of the maximum (print) size of an individual image (200 cm²) and the maximum ground area it represents (e.g. the island of Ireland has an area ~84000 km²). The latter restriction varies by product, so that, for example, a ground area **no bigger** than 50 km² can be published for Land-Form PANORAMA™ data. This generally means that any map larger than ~A5 (200 cm²) is not publishable. It is important

to understand the implications of this restriction. As most peer-reviewed journals are now distributed over the internet, either solely electronically or in tandem with a traditional print process, it means that any OS digital product is effectively unpublishable in any journal. However the ramifications of this policy are more far reaching in that the license also covers all data derived from the original OS data source. Smith (2005a) has outlined an example where JoM could not publish a map that had been accepted for publication due to these restrictions.

DATA DISTRIBUTION

The move towards open access publication has, in part, been driven by the requirements of funding bodies to access research results. This mandate has also been extended to data deposition, with institutional repositories often accepting data. In addition there are also data centres maintained by funding bodies (e.g. Natural Environment Research Council; <http://www.nerc.ac.uk/research/sites/data/>). Despite this move towards data deposition, sharing and distribution, *access* to research data remains *ad hoc*, particularly when research has no direct funding. There is often less impetus to share data and this is partly due to research publications driving the research agenda, not data. And, within geosciences at least, there is little historical context for data sharing.

Researcher's willingness to share data is slowly changing, however the publishing agenda also needs to change. In much the same way that JoM provides an opportunity to publish maps, so there should also be an opportunity to publish the observational, analytical, and interpretive data from which the map was created. This must be citable, with an original, definitive, version stored in an appropriate location. And in the same way that journal articles are peer reviewed, so data should also be.

There are further complications to data sharing that remain active areas of research (e.g. GRADE Project: <http://edina.ac.uk/projects/grade/index.html>). These include:

1. Data Sharing: which digital data formats should be used to distribute data?
2. Data Preservation: which digital data formats will allow maximum preservation (durability)?
3. Data Presentation: how can cartographic presentation be preserved?

In terms of data sharing, geospatial data types simplify to raster, vector, and attribute data. These form the lowest common denominator meaning that they should be both easy to share and easy to preserve, however there remains no preservation of the cartographic presentation. Increasingly the complexity of data formats will allow a greater richness in data sets and, in particular, data presentation (e.g. an ESRI .mxd file), however this will often be at the cost of proprietary formats and therefore the ability to both share and preserve data. Simple, open, data formats remain an essential starting point. At JoM, data can be incorporated as part of the journal article and is distributed in the following formats:

1. **Vector:** the Shapefile remains a good file format as it is openly published, with open source tools available to manipulate them. Unfortunately it does not preserve topological relationships and a format such as E00 might be required in these situations.
2. **Raster:** GeoTIFF remains the predominant open format that is widely supported, however both JPG and JPG2000 have wide application and remain suitable for distribution.

3. **Attributes:** there currently remains no open database format ideally suitable to supporting the distribution of attribute data. DBASE IV is widely supported, however text-base CSV remains the lowest common denominator.

One format that isn't noted above is the Geographical Markup Language, an Open Geospatial Consortium specification for the encoding of geospatial data (<http://www.opengeospatial.org/standards/gml>). Whilst only ratified in 2002, over 40 vendors formally support the format and it has the potential to provide a widely supported, open format for data distribution.

JoM supports the distribution of data and, in particular, the ethos that "data should follow the research." Separating data and content is problematic in that an explicit link needs to be maintained between the two. Journals from the 19th century were able to distribute "data" in the form of photos, tables, maps and drawings and it is clear how effective this immediacy was. The peer review process also becomes much simpler.

The distribution of data also touches upon both open access and copyright. With an increase in OA journals the potential to have free and immediate data distribution is an attractive proposal. However there needs to be considerable care taken in resolving remaining copyright issues. Not only are appropriate licensing arrangements for third party data required, but the vested interest of stakeholders in data that is deposited must be considered. Third party copyright remains an active research area, although initial reports suggest that this may not be as restrictive as previously thought (Waelde and McGinley, 2007). The issues related to stakeholder interests, however, requires further investigation. For example, what are the interests of the employer, employee, funding agency, and any co-investigators?

Additionally, the complexity of dealing with data formats and, in particular, data volumes, makes coupling data distribution and publishing a difficult task. The *de facto* solution is for an author to reference their data as residing in a separate data repository, although many request that third parties contact them directly. However there is a strong argument to maintain the article-data link, placing the emphasis for data review, storage and dissemination with the publishers. This may increase the cost of publishing, but would provide a consistent approach, across disciplines.

CONCLUSIONS

This article has covered material related to journal publication, yet the central thrust is actually far more general. Grant awarding bodies do not want to pay to view research they have already funded and the preference is that these are "free to access" (i.e. open access). OA does not necessarily mean a "traditional" journal, but can also incorporate e-only journals, institutional repositories, subject repositories, and grant body repositories. These may or may not incorporate peer review as part of the lodgment procedure. Preservation of research remains a central issue and it is clear that publishers take this issue seriously with, for example, Elsevier and the National Library of the Netherlands (Elsevier, 2007) providing a permanent digital archive for current journals. All of the content published by JoM is permanently archived by the British Library under the UK Web Preservation Programme (<http://www.webarchive.org.uk/>).

Preservation goes beyond the publication of research results in journals, with the requirement in the United Kingdom for research grant holders to deposit research data in

repositories. Data distribution is a key component in any data preservation strategy and should be closely linked to the published, peer reviewed, research outputs. For the time being, journal publication and data repositories remain separate entities, however there is a strong argument to, at least in part, merge the activities of these two areas.

There also remain more practical considerations, with a move to greater emphasis on data distribution. In particular, what data formats should be used for the storage of spatial data? These need to be fit-for-purpose and open, with a high potential for preservation. Geospatial repositories remain in their infancy and the subject of active research, such as the GRADE Project (JISC, 2005).

The final link in this chain revolves around legal aspects of data sharing and, in particular, copyright. The above discussion outlines some of the fundamental issues concerning the use and re-use of research outputs and geospatial data. Can the interests of multiple stakeholders be satisfactorily resolved to allow data distribution? Should geospatial data be considered under copyright or, in Europe, is the Database Directive (covering “collections” of data) more appropriate (Waelde and McGinley, 2007)? Are licensing restrictions concerning the use of derivative datasets satisfactory? Whilst the first question remains unresolved, the second has been highlighted as being actively researched. Derivative datasets are an important area, with licensing conditions varying between suppliers. For example, Intermap, with the sale of the NEXTMap Britain product, does not claim intellectual property rights for thematically derived data (Intermap, 2005), whilst the OS do claim these rights (EDINA, 2007).

ACKNOWLEDGEMENTS

Some of this material appeared in an earlier article (Smith, 2005b) and is now significantly updated.

REFERENCES

Biomed Central, 2004, What is Biomed Central? Accessed at <http://www.biomedcentral.com/info/>.

Creative Commons, 2005, Learn More about Creative Commons. Accessed at <http://creativecommons.org/learnmore>.

Copyright Office, 2007, Copyright Law of the United States of America §105. Accessed at <http://www.copyright.gov/title17/92chap1.html>.

*

Delamothe, T., and Smith, R., 2003, Editorial: British Medical Journal, v. 327, no. 7409, p. 241-242.

EDINA, 2007, Digimap: Ordnance Survey Data Sub-licence Agreement. Accessed at <http://www.edina.ac.uk/digimap/ostterms.html>.

DRAFT -- To be published in DMT'07 Proceedings
(see <http://ngmdb.usgs.gov/Info/dmt/>)

Elsevier, 2007, Permanent digital archive assures perpetual accessibility of scientific heritage. Accessed at http://www.elsevier.com/wps/find/authored_newsitem.cws_home/companynews05_00020.

HMSO, 2004, Scientific Publications: Free for all? Science and Technology Committee, tenth report of Session 2003-04: Stationery Office, London.

Intermap, 2005, Product Handbook. Accessed at <http://www.intermap.com/images/handbook/producthandbook.pdf>.

JISC, 2005, Scoping a Geospatial Repository for Academic Deposit and Extraction. Accessed at http://www.jisc.ac.uk/index.cfm?name=project_grade.

Morris, S., and Powell, A., 2005, Dissemination of and access to UK research outputs: response from the Association of Learned and Professional Society Publishers (ALPSP) to the RCUK position paper, accessed at <http://www.alpsp.org/RCUKResponse.pdf>.

Research Councils UK, 2005, RCUK Position Statement on Access to Research Outputs: Research Councils UK Report, p. 2, accessed at <http://www.rcuk.ac.uk/access/default.htm>.

Rightscom, 2005, Business models for journal content. Final report to Joint Information Services Committee, p. 64, accessed at http://www.nesli2.ac.uk/JBM_o_20050401Final_report_redacted_for_publication.pdf.

Rowland, F., 2002, The peer-review process: Learned Publishing, v. 15, p. 247-258.

Smith, M. J., 2005a, The Journal of Maps: an electronic journal for the presentation and dissemination of map based data: Journal of Maps, v. 2005, p. 1-6.

Smith, M.J., 2005b, Open access journal publication: methods of implementation and copyright issues: Society of Cartographers Bulletin, v. 39, p. 21-24.

Waelde, C., and McGinley, M., 2007, Designing a licensing strategy for sharing and re-use of geospatial data in the academic sector. Final report to the GRADE Project, p. 40, accessed at <http://edina.ac.uk/projects/grade/gradeDigitalRightsIssues.pdf>.