

Defining and Advocating Open Data in Archaeology

Stefano Costa

Working Group on Open Data in Archaeology, Open Knowledge Foundation

Anthony Beck

University of Leeds, United Kingdom

Andrew Bevan

University College London, United Kingdom

Jessica Ogden

L – P : Archaeology, United Kingdom

Abstract:

“A piece of content or data is open if anyone is free to use, reuse, and redistribute it - subject only, at most, to the requirement to attribute and share-alike” (<http://opendefinition.org/>). Driven by demands for greater transparency from government, general freedom of information and an increased awareness of the unanticipated re-use values of existing information, Open Data has seen dramatic growth in the past two years. Is archaeology part of this general trend? Our aim is to explore what it means to make archaeological data open and what processes are required to make it happen in a satisfactory way. There are three major goals: (a) individual and institutional advocacy, (b) ethical discussion and consensus-building, and (c) knowledge transfer (licenses guidance, wider academic context, repositories etc.). In this paper, we explore some of these issues in greater detail.

Keywords:

Open Access, Open Data, Open Archaeology, Open Knowledge Foundation, Linked Open Data, Knowledge Transfer

1. Introduction

‘Open’ is increasingly invoked as an attractive thing to be, across a host of human behavioural domains from scientific practice to corporate responsibility to governmental action (to name but a few). The concept readily suggests notions of public accountability, transparency of practice, plurality of opinion and scientific repeatability, thus providing a catchy banner for the advocacy of a range of perceived public goods. In contrast, while being ‘closed’ might occasionally carry positive connotations of increased security, it is less marketable and causes significant problems for a domain whose advances are predicated on a developing and accessible corpus. Some aspects of the open agenda are now several decades old and have had a tangible impact on many different sectors. For example, two well-established international initiatives are Open Access

(e.g. where publications such as journal articles are made publicly available without a fee) and Free/Libre and Open Source Software (FOSS, where source code is made available under liberal licenses that encourage onward use and modification), both arguably expressions of a wider change in the free circulation of information and knowledge. A small minority of archaeologists have also been embracing, discussing and/or contributing to such initiatives for some time, particularly those with an interest in customising digital tools or encouraging broader archaeological participation and dialogue beyond a few well-resourced Western institutions (e.g. in line with the participatory agenda espoused by World Archaeology). Needless to say, the Web has been a catalyst for these changes.

The latest chapter is Open Data, where priority is placed on full publication of vast tracts of undigested information that previously might only be referred to publicly in summary form, if

Corresponding author: stefano.costa@okfn.org

at all. As with Open Access and FOSS, Open Data is not an initiative that is specific to archaeology, but instead reflects much wider calls, for example, for increased access to government data (data.gov.uk, 2012) and data across academia (Royal Society, 2012). The underlying rationale of Open Data is that promoting untrammled access to large amounts of 'raw' information enables patterns of re-use and knowledge creation that were previously impossible and/or largely unanticipated. We seek to define what we mean by archaeological data in clearer terms below, but regardless of how this is done, most practitioners would certainly agree that archaeology produces a great deal of data and that a rapidly increasing portion of this is digital. Indeed, being digital is a necessary condition behind the free circulation of data over the Internet, and it is unsurprising that Open Data is particularly popular in the IT domain. In recent years, this popularity has generated an interest in Open Data in archaeology which is highlighted by the increasing number of sessions at the Computer Applications in Archaeology meetings dedicated to this topic. While there is undoubtedly an element of hype, this growing enthusiasm remains a very positive trend in our view. In this paper, we therefore explore how we might define the nature of archaeological data, and thereafter use this as a backdrop for discussing the strengths and weaknesses offered by existing models for data sharing. In particular, we are interested in advocating data-sharing models that emphasise sustainability, remain aware of different cultures of practice and that engage with broader social and ethical issues in archaeology.

2. What is Archaeological Data?

A traditional, often implicit view of archaeological 'data' places it in opposition to the act of 'interpretation'. From this perspective, data are (often standardised) information packets that are meant to capture the archaeological record or other kinds of archaeological evidence in raw form and which are typically obtained via a rigorous, scientific method (e.g. through controlled excavation or through instrumental analyses), whereas interpretation is seen as a subsequent, richer stage of the archaeological process. Wider debate in archaeology as a discipline (e.g. between so-called processual and post-processual perspectives) and value judgements about more and less important

kinds of archaeology (e.g. digging versus synthesis and theorising) have often found purchase around this assumed contrast. In recent years, however, this contrast has rightly been subject to revision, as various commentators have noted the widely acknowledged, if often forgotten, fact that all stages of archaeological practice involve theory-laden assumptions, and hence that data collection and interpretation are closely entwined. Current perspectives on this kind of issue tend to emphasise pragmatism in making such distinctions (e.g. Lucas 2012), and in line with this, here we retain an idea of data as being packets of information that are particularly amenable to easy reorganisation and/or automated processing. Even so, the kinds of information that can be thought of as data are much more diverse than some might initially think. Through interpretation and knowledge acquisition, data can be transformed. Sometimes the data transformation process consists of a series of steps, or a workflow. Each step can contain assumptions and generalisations that have an impact on the final interpretation. It is rare for this transformational process to be documented. Open archaeological data therefore should refer to all archaeological information that is shared (by whatever means) under licenses that allow re-use by everyone (e.g. according to the Open Knowledge Definition²). This definition is deliberately not a neutral description of the status quo in archaeology, but rather aims to provoke a wider shift in perspective about what constitutes archaeological data in the first place. We should also expand our sense of who the stakeholder might be for open data, avoiding the idea of an experts-only environment to one in which the re-users might be archaeologists, policy makers, planners and the public. Put simply, the data requires democratising. We hope the knock-on effect of such an effort will also raise digital awareness and literacy. This will affect the creation, management and use of data at all levels.

A broader definition of data enables a richer environment of reuse to develop. For example, through the removal of paywalls Open Access results in increased public availability of journals and publications in digital format. However, a fully open license also allows us to add value to Open Access texts by using powerful text-mining software

² Open Knowledge Definition <http://opendefinition.org/>

in order to perform automated categorisation and extraction of information. This is already popular in biomedical science (see Open Knowledge Working Group on Open Science 2012). By means of text-mining, information that was previously thought to be locked within the semantic complexity of natural language can be reborn and extracted as structured data. In line with this, the concept of a 'derived dataset' is key to the understanding of critical aspects of usage licenses and the determination of rights ownership, discussed below, but also for a long-term vision of how archaeological knowledge might circulate if made free to do so. Certain kinds of open archaeological data have the potential to be 'big data', not least because they can claim to have wide spatio-temporal coverage. However, it would be wrong to assume that only big data deserve to be open. As with anything else, scale and size should not be the sole parameter to assess quality. Models of dissemination may differ significantly depending on the size of the dataset, but small-scale datasets are both important in their own right, and prone to becoming large-scale via judicious aggregation with others at a later date. These initiatives allow us to rediscover a pathway to the kinds of synthetic perspectives on long term human culture last possible many decades ago (Bevan 2010).

3. Models of (Open) Data Sharing

It is clear that the sharing of archaeological data, especially via the Web and via licenses that encourage re-use, is nothing new or pioneering. Well-known and established services such as the UK Archaeology Data Service³, Open Context⁴ and the Digital Archaeological Record⁵ host a range of archaeological datasets. Private bodies such as Oxford Archaeology make their grey literature available as an open archive⁶. Research projects share through their websites parts or all of the data archives created during the project. Occasionally, researchers integrate their own published papers with digital data otherwise unavailable through traditional means. Some research teams have their

³ Archaeology Data Service <http://archaeologydataservice.ac.uk/>.

⁴ Open Context <http://opencontext.org/>.

⁵ the Digital Archaeological Record <http://www.tdar.org>.

⁶ Oxford Archaeology Library <http://library.thehumanjourney.net/>.

archives on the Web, but under restricted access. The variety of actors engaged in data sharing is largely mirrored by the variety of technical choices and tools that are adopted to fulfill the task. The spectrum ranges from interactive web applications (through WebGIS or query-able databases) to catalogues conforming to technical standards (such as the European Commission INSPIRE Directive⁷), and from complex datasets made interoperable via simplified formats and structures) to simple datasets available as web pages or documents. It is still rare to publish datasets alongside finished articles and rarer still to have any means to reproduce analysis and processing chains as part of the standard review process.

Given this wider context, we can make a distinction between two alternative models for sharing data: (a) 'living' datasets that either involve interactive applications or continue to experience incremental updates, and (b) static datasets released as a stable resource (i.e. similar to a paper that does not undergo any changes once published). These two models are not mutually exclusive, and technically it is possible to create snapshots from active datasets, or to develop interactive applications for the exploration of inactive datasets.

Some of these web applications (the first model) are easy to install and quite popular because of their ability to perform queries and explore map data via an intuitive user interface. However, this dynamic model is not without issues. First, the scalability of these systems is comparatively low, both in terms of long-term management and usability. Second, their duplicability is low as well, making long-term preservation problematic at best: such server driven systems tend to become one-offs. Moreover, in many cases, actual datasets are locked up in the web service and cannot be easily downloaded. The risk is that when the web service is eventually deprecated the data can be lost. Third, no matter how flexible the design of the interface and querying system, there will always be a limit to what users can do (i.e. the range of possible avenues for re-use is heavily constrained). Finally, web services in general are more difficult to develop and maintain as fully open systems (e.g. as discussed in detail in

⁷ European Commission INSPIRE Directive <http://inspire.jrc.ec.europa.eu/>.

documents like the Franklin Street Statement⁸ and the Open Software Service Definition⁹).

The second model, involving static datasets, is simpler from a technical point of view, and allows for the sharing of files and file archives by users who download them from, for example, a stable URL. The absence of built-in exploratory tools makes documentation particularly important in this case, even though catalogue software like CKAN are developing such tools. Documentation need not be limited to a formal description of the content, but should also be explicit about the who, when, where, what, why and how the dataset was created in first place. Data can be published in various formats, with some interoperable formats which have demonstrated widespread acceptance (.csv, .xml, .shp). However, the use of some of these interoperable formats can minimise the complexity of the underlying source data leading to the publication of incomplete datasets. In terms of data 'objects' this model is far more scalable, as from a technical point of view archiving is straightforward, and the cost of hosting even thousands of these datasets is linear to the quantity, not the quality or complexity. However, data aggregation for analysis beyond that envisaged by the original scale of data collection is complicated due to structural, syntactic and semantic heterogeneities (Bishr 1998).

In reality, the main data repositories cited above allow for some provision of both living and static data. Nevertheless, it is worth highlighting the strong analogies between the static model for data dissemination and traditional publication (either digital or printed). Hence, static approaches to data sharing not only are attractive for the technical reasons noted above, but also because they encourage a smoother transition from older (static, printed matter with restrictive licensing) to newer (static open data) knowledge sharing systems. Moreover, this approach can also work well within the current system of academic evaluation and reward, by making the dataset a typical type of publication such as a formal data paper. There are already some interesting examples in this direction,

such as the *Journal of Open Archaeology Data*¹⁰ which offers a concept of metadata that is not merely technical but also includes an explanation of the value of the dataset and the possible use-cases envisaged by the authors. The dataset as a finished and fixed item also represents a point of confidence for those who cite or use it, as happens with public domain radiocarbon calibration data.

A third emerging model is Linked Open Data (LOD), discussed only briefly here. A fundamental aspect of LOD is the use of Uniform Resource Identifiers (URIs), unique references that can be associated with both human- and machine-readable resources. LOD datasets are commonly distributed as downloadable, highly interoperable files (usually plain text). In one sense, LOD is quite similar to the static data model outlined above, but the use of URIs tends to encourage their treatment as a physical node in a wider network accessed via accompanying web service. Advanced querying of these kinds of LOD frameworks is possible via the SPARQL protocol, in a way which ushers in a highly interactive data environment, albeit one that at present can only be provided by well-resourced institutions. LOD requires datasets to be harmonised to some degree but this need not be very onerous for dataset creators. For example, tools such as Google Refine, are a helpful way to clean up data with an eye to LOD resources (such as those provided by Freebase). Even so, we would stress that these more involved tasks, associated with the construction of richer semantic web content, can and should be kept separate from the easier one of publishing open data in simple open formats. If we continue to place an emphasis on simple open data, we leave fewer excuses for not participating and yet also leave the door open if someone else is interested in enhancing the data via LOD at a later date.

4. Ethics

It is also worth discussing the ethical issues surrounding open data here, with regard both to the reasons why people should feel it incumbent on them to produce such data as part of their professional archaeological lives, and with regard to the potential misuse of open data in the archaeological domain. The socio-economic argument is particularly

8 "Franklin Street Statement", last modified July 14, 2008, <http://autonomo.us/2008/07/14/franklin-street-statement/>.

9 "Open Software Service Definition", last modified October 8, 2008, <http://opendefinition.org/software-service/>.

10 Journal of Open Archaeology Data <http://openarchaeologydata.metajnl.com/>.

compelling. The majority of archaeology is carried out in the interest of the public, often with public money and it is therefore only right that the public get access to the objects and outputs of activities that they have funded. This is particularly clear-cut for academics working at publicly funded institutions and/or on publicly funded projects. The issue initially appears more complicated in the case of developer-funded archaeology, where the funding comes potentially from private commercial sources and the client is notionally the developer. However, in the UK, for example, The National Planning Policy Framework (NPPF: Department for Communities and Local Government 2012) which describes the planning policies on the conservation of the historic environment in England and Wales stipulates that local planning authorities should 'make this evidence (and any archive generated) publicly accessible'. Hence, open data should be construed in this case as a crucial part of the mitigation-by-record process, and not as an afterthought.

A second argument in favour of open data touches on a problem of particular relevance in archaeology: unpublished research. There are various reasons for non-publication, but the sheer cost of bringing archaeological research (especially fieldwork) to formal publication is often a key issue. In such cases, we can nevertheless expect that some data exists. While some might argue that publishing data without offering a synthetic overview alongside it makes for a very limited resource, it is undoubtedly better to have access to data than to have nothing at all. For authors, there is at least some recognition of the work done, as is already the case with archaeological reports filed in grey literature archives. The opening up of grey data is, in our view, an economic way to meet the minimal obligations to publish research. This obligation has been proposed by the European Association of Archaeologists (EAA 2009) and by some national organisations (e.g. Research Councils UK; Associazione Nazionale Archeologi 2010), albeit sometimes together with a 5-10 year embargo period. Including raw data, publications as part of these regulations would represent a major step not only for open data *per se*, but for the wider aim of professional organisations. In contrast, the current view is that publication equals a short report, much smaller than the total amount of information resulting from archaeological activity. We need to transform our understanding of what constitutes

full and satisfactory publication. We would argue that greater favour be given to data rather than synthesis, and more generally to decouple these as two different kinds of publication.

Whilst retaining this sense of the importance and urgency of open data initiatives, it is still also worth noting some ethical arguments that might call for a more cautious approach to open data. For example, the open publication of archaeological data may conceivably lead to greater levels of archaeological looting, especially where this involves provision of exact spatial locations of sites. However, this argument is still largely theoretical and in dire need of some formal demonstration (that this does indeed lead to greater looting activity). In certain particularly vulnerable instances, it may be sensible to place restrictions over who has access to the finest-scale coordinates for archaeological finds. However, this has the unfortunate effect of making data available at multiple scales of granularity based on trust frameworks that are at present rather immature (e.g. probably just based on institutional affiliation), and there are further licensing implications that make it hard to consider this as open data in a strict sense. Furthermore, placing an embargo on data has significant and unquantified downstream implications for a whole range of legitimate activities that would make a positive contribution to knowledge.

A third ethical issue is the degree to which opening up data may conceivably perpetuate existing imbalances and injustices amongst archaeological institutions and practitioners. For example, at the institutional scale, the promotion of open data as a research output in its own right risks reaffirming the advantage enjoyed by bigger institutions who have the wherewithal to absorb the costs of making data open. At the individual scale, it may conceivably risk making it easier for more senior collaborators, company heads, etc. to swallow up some of the rights of individual data creators (especially those early in their careers) by taking the main credit for the overall products (e.g. with respect to the individual data outputs of excavators versus a project director), but ultimately we suspect the reverse danger that already exists is more worrying: currently, data access seems to scale with seniority (via who you know and how important they are; see Cella and Palombini 2012). However, as open

data increasingly becomes embedded in the data production process, it is likely that suitable metrics will be generated so that individual contributions to data impact and quality can be generated and retained.

5. Good Practice

The previous three sections have dealt with Open Data models and surrounding issues that are still a matter of significant debate, both within and beyond archaeology. There is a growing platform of good practices for (open) archaeological data, some already covered in existing documents (e.g. the Archaeology Data Service and Digital Antiquity's *Guides to Good Practice*¹¹), and most of these have garnered a sufficient level of acceptance to be considered as common culture. Issues surrounding technical interoperability are perhaps the best known and the key point to stress is a need for the use of open formats that are independent of any specific software platform. If we exclude the most common and simple problem cases (e.g. documents made with office suite software), the critical point is that currently there remain far too many *de facto* proprietary standards (e.g. .dwg) and far too many undocumented formats produced by survey, measurement and analysis tools. In contrast to technical format issues, metadata (structured background data describing a given dataset) is a method of documentation that is not very well known outside of expert domains. This is partly because of the lack of domain-specific standards, and partly because metadata are seen as mainly something to be done for large catalogues, while a single dataset is often not deemed worthy of this kind of treatment. Furthermore, some types of metadata are easier to add or more commonly produced (e.g. author and keywords of a document, bounding box of a spatial dataset), while the mere existence of a metadata structure does not necessarily mean it will be used (e.g. photographic images where this functionality is present but rarely used). Clearly, the specificity of each dataset asks for a detailed description of the methods and conditions with which data has been recorded and structured regardless of the nature of data. So, for example, while a photographic database may have all EXIF metadata about exposure time, even spatial location, it may or may

11 "Guides to Good Practice," accessed September 23, 2012, <http://guides.archaeologydataservice.ac.uk/>.

not have information about the aim of the photographer, the criteria for choosing what to photograph, in what detail and so on.

The legal aspects of open data are arguably more complex than the technical ones, and also something that is presently under-addressed in the teaching of archaeology. Indeed, while choosing an open license for data is actually very simple, many people continue to prefer a restrictive closed license (perhaps because they are uncertain about the nature of potential re-use) or do not see the point of defining an explicit license (e.g. they have not realised that an open license is likely to be very beneficial for the circulation of their work). Finally, further problems can arise if open data need to be combined with proprietary data to create derived datasets. Some common cases of this issue occur with satellite imagery, but the same can happen with a photographic catalogue, or a dendrochronological calibration curve. The derived datasets will be necessarily restricted by the proprietary license, and it will not be possible to share them without breaking the license terms.

Bearing these issues of adoption in mind, there is nonetheless now a solid platform provided by two organisations that have been dealing with the problem of legal aspects of data for some time now, Creative Commons (CC) and Open Data Commons (ODC). There are not many licenses for open data, especially if compared to software licenses. We can make a distinction among three different types of open license: share-alike, attribution and public domain licenses. Share-alike licenses follow the idea of copyleft and of "cascaded" sharing of derived works under the same license (for further discussion, see Murray-Rust 2008; Cheliotis 2009; Stodden 2009). Wikipedia and all Wikimedia projects use a share-alike license, CC-BY-SA 3.0, that is appropriate for text and images. The Open Database License (ODbL) is better suited for factual data¹². Attribution licenses only require that copies and derived works maintain an indication of the original authors. In this case, too, the CC-BY 3.0 is best for creative content and ODC-BY for factual data. Public domain licenses are waivers of any right, including the right to be cited as

12 "Open Database License FAQ," accessed September 23, 2012, <http://opendatacommons.org/faq/licenses/#why-not-use-a-creative-commons-or-freeopen-source-software-license-for-databases>

authors, and are appropriate for instrumental data without a clear creative dimension (e.g. chemical analysis), public domain data (bibliographic data), or big datasets that need to be aggregated in several iterations, making attribution tracking quite difficult to achieve. The two available licenses, CCO and the Public Domain Dedication License, are essentially identical, but the formulation of CCO is more universally interoperable with normative systems, having a fallback declaration in those cases when it is not allowed to waive all rights upon one's own work. Note also that three different licenses in the Creative Commons family are mentioned above, each one with its specific traits. Hence, there is no such thing as a "Creative Commons license" or "Creative Commons 3.0" and users should be guided in how they refer to their preferred license by the service available on the Creative Commons website.

Archaeological data collection is governed by a number of statutory, legislative and professional frameworks. These represent the policy environments under which contractual and academic archaeological works are enacted. This policy has a direct impact on how, and on what terms, archaeological data should be archived and disseminated. Unsurprisingly, many of these frameworks advocate the deposition of data and other resources in publically accessible repositories. For example, in the UK the Department for Communities and Local Government (2012, 32) state that evidence from archaeological works conducted as part of the planning process (and any archive generated) should be publicly accessible. Equally for the academic sector, Research Councils UK state that 'Publicly funded research data are a public good, produced in the public interest, which should be made openly available with as few restrictions as possible in a timely and responsible manner that does not harm intellectual property'. Unfortunately, most policy does not make a distinction between digital and analogue resources and, as digital archaeology is relative immature, many of the established repositories are designed for physical archives (plans, contexts, artefacts, ecofacts and synthetic reports). However, the majority of archaeological data in the future will be collected, analysed, interpreted and published *in-silico* (digitally). It is important that the policy environments, which in principle support publicly accessible archive deposition for analogue archives,

should do the same for digital archives and emphasise mechanisms that encourage re-use.

Finally, it is worth highlighting how important it is to develop a broader understanding and stronger awareness of copyright among all kinds of archaeological practitioners. We envisage the development and increasing importance of lightweight forms of copyright tracking and recording (e.g. by means of metadata and automatic tracking and versioning systems). Currently, authors' rights with regard to archaeological data are more often than not an obstacle to the circulation of knowledge rather than an efficient mechanism to protect one's work. As such, these rights need to be standardised and constrained in terms of their downstream effects, not only on active and future work, but also for existing (currently orphaned) works and grey literature.

6. Conclusions

For some archaeologists, Open Data requires a dramatic change of perspective, the acquisition of new skills and a major shift from curation of software to the curation of data. For others, Open Data may be as simple as doing what they are already doing. In all cases, technical development alone is not enough. Technologies must be accompanied by a proper social framework. Open Data has the potential to be trans-formative which means that it could significantly disrupt the sector. Education about the potential benefits of Open Data and the social implications of any transformation are required to take full advantage of the new possibilities of open archaeological data. Therefore, the link between funding, publication and Open Data is a key area, and we reaffirm the need to encourage national funding agencies, and international ones (private or public) to build in Open Data policies into the requirements of their grants and to check for such a track record in subsequent grants. Open Data needs to be a more relevant part of the archaeological publication, research, management, curation and policy process, and not merely an afterthought.

References

Archaeology Data Service and Digital Antiquity. 2012. "Guides to Good Practice." Accessed September 23. <http://guides.archaeologydataservice.ac.uk/>.

- Associazione Nazionale Archeologi. 2010. "Codice Deontologico." Last modified March 10. <http://www.archeologi.org/public/ANA.CODICE%20DEONTOLOGICO.MOD.pdf>.
- Autonomo.us. 2008. "Franklin Street Statement." Last modified July 14. <http://autonomo.us/2008/07/14/franklin-street-statement/>.
- Bevan, A. 2012. "Spatial methods for analysing large-scale artefact inventories." *Antiquity* 86 (332): 492-506.
- Bishr, Y. 1998. "Overcoming the semantic and other barriers to GIS interoperability." *International Journal of Geographical Information Science* 12, 299-314.
- Cella, E., and A. Palombini. 2012 "The future of the profession of archaeology: the protection of intellectual property rights and free access to data." In *Opening the Past. Archaeological Open Data. MapPapers* 3-II: 78-80.
- Cheliotis, G. 2009. "From open source to open content: Organization, licensing and decision processes in open cultural production." *Decision Support Systems* 47: 229-244.
- data.gov.uk. 2012. Open Data White Paper: Unleashing the potential. Accessed 3 February 2013. <http://data.gov.uk/library/open-data-white-pape>.
- Department for Communities and Local Government. 2012. "National Planning Policy Framework." Last modified March 27. <http://www.communities.gov.uk/publications/planningandbuilding/nppf>.
- European Association of Archaeologists. 2009. "The EAA Code of Practice." Last modified September 19. http://www.e-a-a.org/EAA_Code_of_Practice.pdf.
- Lucas, G. 2012. *Understanding the Archaeological Record*. Cambridge: Cambridge University Press.
- Murray-Rust, P. 2008. "Open data in science." *Serials Review* 34 (1): 52-64.
- Open Knowledge Foundation. 2008. "Open Software Service Definition." Last modified October 8. <http://opendefinition.org/software-service/>.
- Open Knowledge Foundation Working Group on Open Science. 2012. "Working Group Response to IPO Consultation on Text Mining Copyright Exception." Last modified March 21. <http://science.okfn.org/2012/03/21/response-to-ipo-consultation-on-text-mining-copyright-exception/>.
- Royal Society. 2012. *Science as an open enterprise*. London: The Royal Society.
- Stodden, V. 2009. "Enabling reproducible research: licensing for scientific innovation." *International Journal of Communications Law and Policy* 13: 1-25.