

Beyond Open Access: (re)use, impact and the ethos of openness in digital editing

Anna-Maria Sichani

1. Introduction

'Open Access' is something of a buzzword in academia and in digital scholarly editing. My paper aims to critically engage with digital scholarly editing and Open Access through the lens of reuse and value creation. Such an approach distances itself from discussions of funding models needed to cover the costs of providing free access to expensive scholarship as well as from copyright and licensing issues. To be honest, it aims to get one step further: **beyond Open Access** as we used to think of. I will first attempt to describe how (what I call) an 'enlarged ethos of openness' succeed in promoting creative reuse and redistribution of diverse digital content and data and how reuse is linked with and contribute to the value, impact creation and sustainability of digital data. I claim that such an approach of **assessing value through reuse** should be central in contemporary discussions of digital scholarly editing and Open Access, given that the majority of digital editions result from publicly funded projects, where the scholarly content and its value are need to be reconceived and redefined outside the sphere of monetary and material exchange, outside traditional economic/ market thinking.

In my main analysis, I will first try to map (though a qualitative and quantitative approach) the current practices employed in digital editions towards Open Access and reuse. Although the information environment in which digital scholarly content is created and delivered has changed phenomenally over the past fifteen years, allowing the sharing and reuse of digital data, and though the number of publicly accessible digital editions remains on the increase, I argue that limitations in adopting an Open Access agenda focused on reuse in digital scholarly editing still persist. I will then attempt to reveal the main dissuasive reasons for such a stance by questioning mainly the degree to which the patterns of reuse in scholarly editing have changed as we have moved from print to the digital. Finally, I will try to present and further discuss the different models of reuse in digital scholarly editing on both a theoretical and practical level, hopefully laying out a persuasive argument for the multiple benefits of such an endeavour.

2. The virtuous cycle of open digital content: re-use and value

2.1. As much as it gets: Open Access and the 'ethos of openness'

Open Access: what's in a word? Originating in the "free" software movement and soon adopted in the early 2000s by academics and libraries calling for "free immediate access to, and unrestricted reuse" (PLOS, n.d.) of scholarly research, Open Access is grounded in a dual rupture as regards with "price and permission barriers" (Suber 2012). In the past decade, the culture and practices of Open Access have been strategically expanded and reinforced by the rhetoric of Open Definition, celebrating content/ data that is "freely used, modified, and shared with anyone for any purpose" (Open Definition, n.d.). Such an expansion and insistence on defining the "openness" in the digital information environment marks the growth of open-related initiatives and communities of practice (eg. Open Data, Open Source, Open GLAM) while also helping us uncover the major principles of 'openness': 1) Availability and Access, 2) Re-use and Redistribution, 3) Universal Participation.

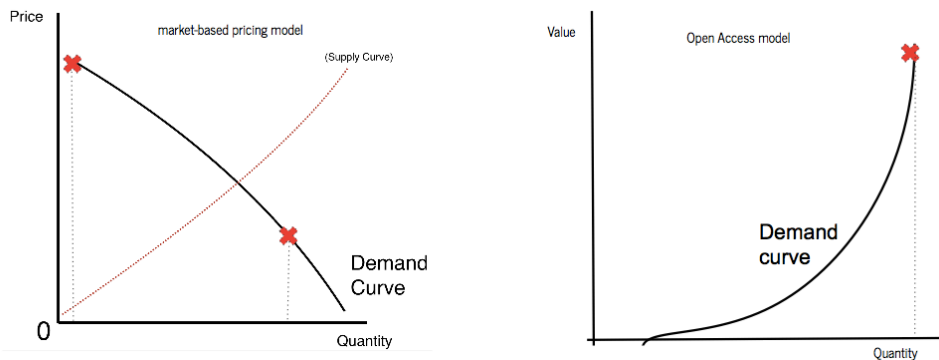
Though the OA movement has made great strides in the last decade, it is important to celebrate all the small and big victories while also understanding that there is a complex matrix behind Open Access, meaning various degrees and combinations of the abovementioned principles. Speaking of barriers, there are two non-equivalent kinds of free online access: to adopt the terminology established for scholarly publications in journals, there is 1) "gratis OA", which removes price barriers but not permission/copyright barriers, and 2) "libre OA", which is free of both price barriers and unnecessary copyright and licensing restrictions, allowing reuse rights which exceed fair use/fair dealing enshrined in copyright law. There is room for variation here, as there is more than one kind of permission barrier to remove and more than one way to do it (type of re-use: copying, redistribution, derivative works, commercial re-use // attribution & right to remove attribution // indication of modification) and this is what Open Licenses (such as [Creative Commons](#) or [Open Source Licenses](#)) are designed to offer, i.e. a simple, standardized way of sharing and reusing works under a choice of conditions. Besides different types of barriers, there is also variety on what is actually open, speaking of output level and source-files and documentation level.

2.2. Open and useable digital content: value through reuse

Adopting such an 'ethos of openness' must necessarily reflect on the existing value systems of both academia and the market: in brief, it is mainly about disaggregating the concepts of value from cost and price as the

conventional law of demand and supply and the scarcity principles teach us. Instead of making content valuable by making it scarce, as it is the rule in traditional economics and cost/market-based pricing, **Open Access makes new knowledge/data valuable by making it widely available and open to reuse.**

Law of demand / supply - Scarcity principle



Such an understanding of **value through (re)use**, centres around the concept of non-rivalrous commodity exchange (Eve 2014; Suber 2012), introducing a remarkable dissonance with traditional evaluation paths for scholarly or cultural content. All in all, **the idea of re-use** could be used **as an alternative metric** with which to assess the value and impact of digital resources.

In light of the expanding mass of digital content created in the last decade, there has also been an ever-growing research interest in the area of use and impact assessment for open digital content, which is an incredibly difficult -if not impossible- task; many initiatives have developed and applied an array of both qualitative (stakeholder interviews, resource surveys, user feedback, focus groups, and questionnaires) and quantitative methods (e.g. webometrics, log file analysis, scientometric [or bibliometric] analysis, and content analysis) (TIDSR, Hughes et al. 2013) that capture “information about the whole cycle of usage and impact” (Meyer et al. 2009, p. 6) in order to assess the usage & impact of digital resources (Warwick et al. 2006; Hughes 2012; Meyer et al. 2009; Tanner 2012).

Though this interest is mainly the result of pressure exerted by funding bodies (eg. NEH, AHRC, Wellcome Trust, AHRC, EU/Horizon 2020, Mellon) and governments insisting on “the need to demonstrate the ‘impact’ of publicly funded resources and research, as a means of quantifying the value of the investment in their creation” (Hughes 2012, 2), alongside guidelines of major funding bodies specifying that projects’ outputs (esp. software) should be “free in every sense of the term, including the use, copying, distribution, and modification” (NEH 2013) and the opening up of many GLAM institutions datasets (such as [DPLA](#), [Reiksmuseum](#), [Cambridge Digital Library](#); see also Terras 2015) for reuse and redistribution, it demonstrates nevertheless that stakeholders are now strategically moving from the creation of digital content towards a more sophisticated awareness of how notions of reuse, value, impact, open access and sustainability remain entwined.

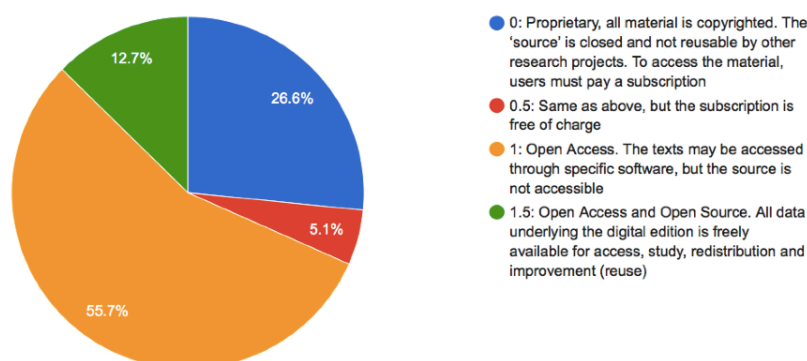
3. Digital Scholarly Editions made (not-so) open

3.1. Reading the numbers

How does this embracing of the virtuous cycle of open digital scholarly and cultural content and such an ethos of openness promoting ‘value through reuse’ relate to the field of digital scholarly editing?

In order to answer this question, I have collected data through qualitative analysis of digital editions projects and I have also extracted some quantitative data from Greta Franzini’s [catalogue of Digital Editions](#) (total: 210 editions), to which I contributed.

Catalogue of Digital Editions and Open Access



Even though, historically, digital editions were conceived and developed with the aspiration of “*making [J]’s vast work freely and conveniently accessible to scholars, students, and general readers*”, what digital editing endeavours actually achieves is more a critical curated digital reunification/gathering of materials scattered around the world in libraries and private collections; in other words, it succeeds in breaking down the textual scholarship barriers of time and space as regards the primary material. In addition, as a vast number of Digital Scholarly Editions are initially developed under grant cycles from funding agencies and governments, they have from the outset to conform to OA in the sense that they need to be free of charge. Despite persistent rhetoric regarding accessibility and availability, a relatively small percentage of those (only 12,7%) make their ‘beyond-the-output content’ openly available for reuse and redistribution.

3.2. Revealing the barriers

Thus barriers still remain, both in the conception and in the practices of what it means for a digital edition to behave in an open access manner. The reasons for this may vary.

First of all, digital editions consist of cross-domain and heterogeneous primary material (eg text transcriptions, facsimiles, critical apparatuses, indices etc.) and thus issues of copyright may still persist or there might be different licensing frameworks for each of these materials (eg. copyrighted facsimiles, pictures from the 20th and 21st c, orphan works etc). Moreover, behind the scenes digital editions usually consist of far more than the output level: encoded XML source files, XML schemas, processing and transformation scripts, entity relation models, rich metadata, transcription and XML encodings, encoding conventions & decisions, project documentation (feasibility studies, proposal) and so on. All of them, are interpretative decisions or editorial choices and integral parts of the project’s workflow.

Usually, as digital editions are developed within project-based frameworks, in which teams and contributors perform flexible roles, and allocated with fluid or overlapping tasks, claim ownership or authorship of digital data is difficult to determine. In addition, a fear remains regarding losing control over your data (meaning that you are no longer able to update it) as well as a fear of malicious (re)use.

Aside from legal, economic, or operational reasons behind licensing complications, I strongly believe that the “page paradigm” (Sahle 2008; Pierazzo 2015) remains a crucial hindrance in adopting an openness ethos that will encourage reuse; its inheritance is still so strong in our scholarly culture that we remain “zoned to print” (Sutherland 2009, 20), thus tending to create and use digital editions “as restrictedly repositories of data and generators of print editions” (Sutherland 2009, 2), as end-products or self-contained entities handed over to the end user “to be seen and not touched” (Shillingsburg 2010; Dahlström 2011, 103). The “look-but-don’t-touch” problem, as Dahlström terms it, is rooted, on the one hand, in a strong scholarly attachment to completed, finished, publishable work, related to enduring issues of attribution of credit and professional reward. On the other hand, the problem is the partial result of a limited knowledge or evidence of how people might interact with and (re)use digital scholarly editions and their components.

4. Re-use’ in the history of modern textual scholarship

I am firmly convinced that we deserve more than a read-only world – and in digital editing as well. Digital scholarly editions remain in an incunabular stage, primarily because we are too reluctant to change our habits when interacting with/ using digital editions, which by definition possess different material qualities from printed ones. In other words, we continue to use them in much the same way as we use(d) a printed edition: for reading, consultation, observation, or in order to incorporate the new findings enabled by the computational technology

into our next printed scholarly output (an article or a book). Such a pattern of reuse differs little from the long established model of reuse in print culture: the footnote, the apparatuses and the related reference systems. While footnote “offer the empirical support for stories told and arguments presented” (Crafton, 1999, vii) and manifests a social model of knowledge creation, it also suggest a fertile but freezed interaction with the source, whose terrain of production and reproduction was defined by the print culture. As “electronic texts [and editions] are artefacts or mechanisms... amenable to material and historical forms of investigation, [thus] challeng[ing] textual critics to respond to the new medium in terms of its own materiality, architecture and functioning, as distinct from those of print” (Sutherland, 2009, 22), I would suggest that the real advantages of digital editing will become apparent as we further understand and reuse digital scholarly editions in terms of their very own materiality and functioning, thus advancing us beyond ‘screen essentialism’ and exploring new patterns of reuse.

5. Towards an ethos of openness and reuse in digital editing

5.1. How we can reuse a digital edition?

There have been several theoretical proposals for how a re-use approach might be applied to digital editing. We can trace a few early but brave examples of such a practice.

In their proposal for *Open Source Critical Editions*, Bodard and Garcia (2009) support the distribution of the raw code, the documentation of the tools and applications that were used in reaching these conclusions, and the methodological statement behind the output, claiming that this is not only a commitment to operational transparency; it also amounts to a critical stance, in the sense that it “implies the adherence to reasonable methods and principles”. Getting “full access to the raw code is what that makes an experiment or a solution reproducible” otherwise “it is a dead end; it cannot be built upon”.

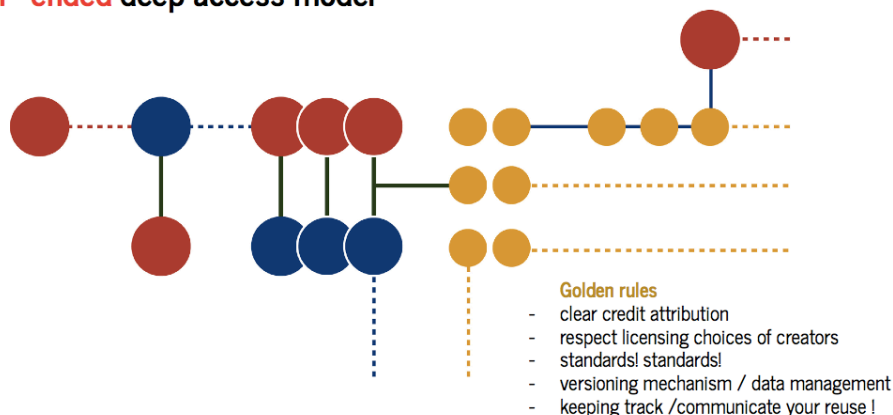
In a similar vein, James Cummings (2009) claims that the model of “agile editions” (easily transformed, repurposable) based on the stand-off encoding proposal presupposes the opening up and distribution of underlying XML files alongside the HTML versions rendered. Moreover, by having first-hand examples of bot successful and unsuccessful forms of community practice, we will be able to discover of our joint errors and misunderstandings of the Guidelines and thus aid into their improvement.

Peter Boot and Van Zudert’ s proposal for “digital scholarly edition 2.0” (2011) suggests a decentralised, diverse and distributed architecture of openly available sources, services and functionalities (virtual research environment), initially enabled through cloud computing and cloud storage. This will reduce the amount of custom developed software and enhance the permanence of our digital outputs.

5.2. Models of reuse and/as aspects of value

Let me further extend the abovementioned abstract models by introducing how **an open-ended model of deep access** can be implemented in digital editing projects and mainly discuss the potential value from such a reuse ethos. Its foundations mainly rely on Dynamic systems development principles as found in the agile project delivery framework. In the open ended deep access model digital scholarly editions tend to be seen not as a static, closed-ended project or fixed product but more as an ongoing and open scholarly enterprise, as a continuum of iterative and incremental reuses of the various components of the project.

open - ended deep access model



Deep access calls for opening up for reuse and redistribution source files (encoded XML, related schemas & ODDs, scripts for processing and visualisation, transformation scripts and query algorithms, entity relational models etc.). By going even deeper, we might also think about sharing parts of the project's documentation such as funding proposals (successful or not), feasibility studies, editorial conventions, even budget plans (as in most cases we are speaking of publicly funded projects) as supportive material for 'how to get a digital editing project off the ground'.

It's not the Open Access rhetoric per se but we need to find more elaborated ways to question, reveal and assess the value and the impact of such an ethos of openness in digital scholarly editing. By using some of the available quantitative methods to measure the impact of a digital edition (such as web analytics or log file analysis), we might gain important data about file requests, but it's extremely doubtful if we are going to have any insight in the actual trajectory and value of its reuse. We need to engage in a more specialised and qualitative analysis of the patterns of reuse.

Aside from guaranteeing transparency, opening up and sharing source files also yields raw material for a new kind of research and scholarship that could be re-integrated and built upon it; an ideal test-bed for new computational/analytical approaches, either separate from or elaborated through a big data approach; an easy way to produce derivatives and varying outputs on demand (ePub, PDF) with potential for commercial exploitation; an opportunity for improvement and refinement; and finally, exemplary reference material for teaching purposes. An excellent example with such open 'fertile files' can be found at the [Jonathan Swift Archive](#), a project that opens up for [download and reuse](#) its source files and scripts (TEI-XML, XSLT) so that researchers in the Centrum voor Teksteditie en Bronnenstudie (KANTL-CTB) have created a prototype with parallel readings of versions of Swift's works.

Reusing openly available source files may also contribute to reduce development costs, making future projects and their funding proposals more competitive. Though there has, until now, been a 'black hole' surrounding the economics of digital editing projects, I think we could easily claim a great potential for cost and time-avoidance: Latest statistics from Transcribe Bentham argue that not only could we save at least £400,000 if the remainder of the Bentham Papers were transcribed by volunteers, but that the *Collected Works of Jeremy Bentham* (which will run to approximately 70 volumes), could save up to 6 months of research staff time per volume (Causer et al. 2016). Additionally, the reuse of existing XML files contributed to a considerable reduction in the overall budget estimation, making more competitive and finally successful the proposal of the new AHRC funded Bentham project.

Finally, re-use and re-integration of existing data may also provide a distributed archiving solution ([LOCKSS](#), [Tapas](#)), as well as a sustainability venue for continuous refinement and update, securing its long-term sustainability. By adopting such an approach familiar from the sustainable environmental management agenda: while by reducing new demand for raw products, we succeed to avoid depletion, to exploit the maximum of existing resources and thus to ensure the longevity of the planet, in a similar vein

6. Conclusion

To conclude, I think it is a right-timing moment to revisit Open Access in digital editing by focusing and further developing aspects of a more radical approach towards sharing, distributing and reintegrating digital content and outputs and the value of such an undertaking. I strongly support that is not enough to claim for Open Access per se and struggling on the economical & distribution aspect but it would be more important to elaborate, develop and adopt practices and models that will help us to have more Open, Usable, Used and Useful digital scholarly editing future.-

Selective Bibliography

- Bodard, Gabriel, Juan Garcés, Marilyn Deegan, and Kathryn Sutherland (2009), "Open Source Critical Editions: A Rationale." In *Text Editing, Print, and the Digital World*, 83–98. Digital Research in the Arts and Humanities. Ashgate.
- Boot, Peter, and Joris van Zundert (2011), "The Digital Edition 2.0 and The Digital Library: Services, Not Resources." In *Digitale Edition Und Forschungsbibliothek, Beiträge Der Fachtagung Im Philosophicum Der Universität Mainz Am 13. Und 14. Januar 2011*, 141–52.
- Causer, Tim, Grint, Kris, Sichani, Anna-Maria and Terras, Melissa (2016), " 'Making such bargain': Transcribe Bentham and the quality of cost- effectiveness of crowdsourced transcription" (forthcoming)
- Crafton, Antony (1999), *The Footnote. A Curious History*. Harvard University Press.

- Cummings, James (2009), "Converting Saint Paul: A New TEI P5 Edition of The Conversion of Saint Paul Using Stand-off Methodology." *Literary and Linguistic Computing* 24 (3): 307–17.
- Dahlström, Mats (2011), "Editing Libraries", *Bibliothek und Wissenschaft*, 44, pp. 91–106.
- Eve, Martin Paul (2014), *Open Access and the Humanities, Contexts, Controversies and the Future*, Cambridge University Press.
- Hughes, Lorna M, (2012), "Introduction: the value, use and impact of digital collections". In: *Evaluating and Measuring the Value, Use and Impact of Digital Collections*, Lorna Hughes (ed), Facet Publishing, pp.1-12.
- Hughes, Lorna M. , Paul S. Ell, Gareth A. G. Knight, Milena Dobрева (2013), "Assessing and measuring impact of a digital collection in the humanities: An analysis of the SPHERE (Stormont Parliamentary Hansards: Embedded in Research and Education) Project", *Digital Scholarship in the Humanities*, Volume 30, Issue 2, p. 183 – 198
- Meyer, E., K. Eccles, M. Thelwall and C. Madsen (2009), *Final Report to JISC on the Usage and Impact Study of JISC-funded Phase 1 Digitisation Projects & the Toolkit for the Impact of Digitised Resources (TIDSR)* http://microsites.oii.ox.ac.uk/tidsr/system/files/TIDSR_FinalReport_20July2009.pdf
- NEH (2013), Office of Digital Humanities, Digital Humanities Implementation Grants <http://www.neh.gov/grants/odh/digital-humanities-implementation-grants>
- Open Definition (n.d.), <http://opendefinition.org/>
- Pierazzo, Elena (2015), *Digital Scholarly Editing. Theories, Models and Methods*, Ashgate Publishing, Ltd.
- PLOS (Public Library of Science) (n.d.), <https://www.plos.org/open-access/>
- Sahle, Patrick (2008), *A Catalogue of Digital Scholarly Editions*, available at <http://www.digitale-edition.de/vlet-about.html>
- Shillingsburg, Peter (2010), "How literary works exist: implied, represented, and interpreted". In: *Text and genre in reconstruction: effects of digitalization on ideas, behaviours, products and institutions*, W. McCarty (ed), Open Book Publishers, available at <http://books.openedition.org/obp/658>
- Suber, Peter (2012), *Open Access*, MIT Press.
- Sutherland, Kathryn (2009), "Being critical: paper-based editing and the digital environment". In: *Text Editing, Print and the Digital World*, Marilyn Deegan & Kathryn Sutherland (eds), Ashgate Publishing, Ltd, pp.13-26.
- Tanner, Simon (2012), *Measuring the Impact of Digital Resources: The Balanced Value Impact Model*. King's College London, www.kdcs.kcl.ac.uk/innovation/impact.html
- Terras, Melissa (2015), "Opening Access to collections: the making and using of open digitised cultural content", *Online Information Review*, 39: 5, pp. 733 - 752.
- Toolkit for the Impact of Digitised Scholarly Resources (TIDSR)*. <http://microsites.oii.ox.ac.uk/tidsr/>
- Warwick, C., et al. (2006), *The LAIRAH Project: Log Analysis of Digital Resources in the Arts and Humanities. Final Report to the Arts and Humanities Research Council*. School of Library, Archive and Information Studies. University College London: London <http://www.ucl.ac.uk/infostudies/claire-warwick/publications/LAIRAHreport.pdf>