

Chapter 5

Open Source Critical Editions: A Rationale

Gabriel Bodard and Juan Garcés

In 2006 a group of scholars within the Digital Classicist¹ community began to meet, first electronically and then physically, to discuss a range of issues and strategies that they dubbed ‘Open Source Critical Editions’.² An Open Source Critical Editions (OSCE) workshop was held on 22 September 2006 at the Centre for Computing in the Humanities, King’s College London, under the auspices of the AHRC ICT Methods Network. It was also supported in part by the Perseus Project and the Digital Classicist. This workshop was set up with the aim of exploring the possibilities of, requirements for and repercussions of a new generation of digital critical editions of Greek and Latin texts where the underlying code is made available under an open licence such as Creative Commons, General Public Licence (GPL) or Apache.³ It is our assumption, and our assertion, that these issues and protocols will apply to all humanities disciplines that deal with the publication of critical texts. Although we are a relatively homogeneous group, approaching these questions principally as Classicists, it would obviously be shortsighted and counter-productive to imagine that we do not share a great many of these concerns with scholars from other disciplines who are working toward similar protocols.

Technological questions discussed at this event included: the status of open critical editions within a repository or distributed collection of texts; the need for and requirements of a registry to bind together and provide referencing mechanisms for such texts (the Canonical Text Services protocols being an obvious candidate for such a function);⁴ the authoritative status of this class of edition, whether edited by a single scholar or collaboratively; the role of e-Science and Grid applications in the

1 The Digital Classicist website, wiki and discussion list (sponsored by the Centre for Computing in the Humanities, King’s College London and the Stoa Consortium, University of Kentucky) can all be found at <<http://www.digitalclassicist.org/>>. All urls valid at the time of writing, September 2007.

2 See the full Methods Network report at <<http://www.methodsnetwork.ac.uk/activities/act9report.html>>; contributing papers were given by (in addition to the authors): Sayeed Choudhury, Gregory Crane, Daniel Decker, Stuart Dunn, Brian Fuchs, Charlotte Roueché, Ross Scaife and Neel Smith.

3 Creative Commons licensing, see <<http://www.creativecommons.org/>>; GNU General Public Licence, see <<http://www.gnu.org/copyleft/gpl.html>>; Apache Licence, see <<http://www.apache.org/licenses/>>.

4 Canonical Text Services protocol, see <<http://chs75.harvard.edu/projects/diginc/techpub/cts>>.

creation and delivery of editions. Legal issues largely revolved around the question of copyright and licensing: what status should the data behind digital critical editions have? This group assumed that source texts should be both Open Source (with respect to the editions) and public domain (for the texts themselves), but the specifics remain to be discussed. Attribution of scholarship is clearly desirable, but the automatic granting of permission to modify and build upon scholarly work is also essential. There were also questions regarding the Classical texts upon which such editions are based: what is the copyright status of a recently published critical edition of a text or manuscript that the editor of a new edition needs to incorporate? Administrative questions posed by open critical editions included: issues of workflow, collaboration and editorial oversight (on which the examples of the Suda Online and Pleiades – to name only two prominent Classics projects with experience in this area – and of large projects like Citizendium will provide useful terms of reference);⁵ protocols for publication and re-use of source data. Issues of peer review and both pre and post-publication validation of scholarship were also discussed.

Many of the arguments presented in this chapter are not new or especially startling. Our aim is to collect those elements of scholarly thinking that have some bearing on digital publication into a coherent picture that helps to define the particular assumptions within which we are working. These assumptions and arguments clearly draw upon the work of scholars both within and without the Classics,⁶ and are almost always informed by the contributions of the participants in the OSCE workshop. The particular case made here is of course the work of the authors alone and should not be read as representing the opinions or arguments of any other scholar.

The interests and rationale of this active community, while relatively coherent and implicitly well understood internally, have never been fully documented. This chapter, informed by the discussions of the OSCE group yet representing the particular views of the authors, will focus on three of the core issues, all expressed in the title: (1) the sense and implications of the Open Source model; (2) the connotations of ‘critical’ in this context; (3) the issue of what kinds of edition should be included in such a project – literary, eclectic or individual manuscripts – and what this means for the technologies and protocols adopted. Our proposal is that Classical scholarship should recognize OSCEs as a deeper, richer and potentially different kind of publication from printed editions of texts, or even

5 Suda Online editorial policies, see <<http://www.stoa.org/sol/policy.shtml>>; Pleiades editorial workflow, see <<http://icon.stoa.org/trac/pleiades/wiki/WorkFlow>>; Citizendium policies, see <<http://www.citizendium.org/about.html>>.

6 See, for example, P. Robinson, ‘Current Issues in Making Digital Editions of Medieval Texts – or, Do Electronic Scholarly Editions Have a Future?’, *Digital Medievalist*, 1/1 (2005). Available online at <<http://www.digitalmedievalist.org/article.cfm?RecID=6>>. G. Crane, D. Bamman and A. Babeu, ‘ePhilology: When the Books Talk to their Readers’, in R. Siemens and S. Schreibman (eds), *Blackwell Companion to Digital Literary Studies* (Oxford: Blackwell Publishing Limited, 2007), pp. 29–64.

from digitized and open content online editions. OSCEs are more than merely the final representations of finished work; in their essence they involve the distribution of raw data, of scholarly tradition, of decision-making processes, and of the tools and applications that were used in reaching these conclusions. The protocols and technologies for this manner of publication need to be made available and comprehensible to all textual scholars if the unique advantages and opportunities afforded by digital scholarship are to become entrenched in Classical and other philological disciplines.

Open Source

The use of the term ‘Open Source’ in this discussion is perhaps deliberately provocative: the term is generally used in the context of software engineering projects and collaboratively authored source code. It might be argued that a term like ‘Open Access’ or ‘Open Content’ would be more appropriate to a project involving critical texts rather than algorithms. Nevertheless, we shall argue that the principles of the Open Source movement are basically those of scholarly publication, which traditionally requires full documentation of sources, references and arguments, and allows – nay demands – the re-use of these sources and reference to previous editions in future publications on the same topic.

The origins of the concept of Open Source lie in the free software movement, with operating systems like Unix and GNU, web browsing tools like Mozilla, and the large number of projects that circulate code through Open Source development sites such as SourceForge.⁷ The rationale behind the Open Source movement is not that software should be free merely in the financial sense, so that it should be available at no cost, but that it should be open in the sense of free to distribute, learn from, modify and re-distribute. In other words, Open Source software is not necessarily non-commercial (although it often is), but it is software that is distributed along with its source code and with an explicit licence allowing others to modify, fix or enhance this code and circulate the improved version of the software, along with the code, under the same open licence. (Most Open Source licences prohibit the distribution of derivative software without the source code also being made available.) In the words of the GNU GPL:

When we speak of free software, we are referring to freedom, not price. Our General Public Licenses are designed to make sure that you have the freedom to distribute copies of free software (and charge for this service if you wish), that you receive source code or can get it if you want it, that you can change the software or use pieces of it in new free programs; and that you know you can do these things.

⁷ Cf. n. 2, above; about SourceForge, see <<http://sourceforge.net/docs/about>>; see also Open Source Definition (from Linux Information Project) at <http://www.lininfo.org/open_source.html>; Open Source Initiative at <<http://www.opensource.org/docs/osd>>.

To protect your rights, we need to make restrictions that forbid anyone to deny you these rights or to ask you to surrender the rights. These restrictions translate to certain responsibilities for you if you distribute copies of the software, or if you modify it.

For example, if you distribute copies of such a program, whether gratis or for a fee, you must give the recipients all the rights that you have. You must make sure that they, too, receive or can get the source code. And you must show them these terms so they know their rights.⁸

Software licensed under an Open Source agreement is not therefore wholly in the public domain: it is copyrighted, it belongs to the author, the author's rights are protected, the author will always get credit for it and there are limits on what can be done with it. But the author has prospectively given permission for the content to be reproduced under certain conditions, which vary according to the licence chosen. Open Source is not so much a business model with exclusively economic implications as a strategy based on the belief that cultural advances are made by building upon the creations and publications of those who came before us. Without full access to the raw code, the documentation and the methodological statement that makes an experiment or a solution reproducible, a given publication is a dead end; it cannot be built upon. Certain types of creation are not protected by strong copyright and patents: they are stifled by it. An Open Source licence on a software package or suite does, it is true, make it harder for a single entity to make money from a monopoly on that product, but it is equally true that in the best-case scenario it makes possible collaborative work on a scale never seen before in the programming world. It is possible for several – or several hundred – coders to work on the same problem, to take one another's scripts and improve them, build upon them, modify them, sideline or deprecate them if they are superfluous. These co-workers need never have met nor communicated with one another; they may not have the same interests or goals; but they are all, for a time, working toward the same ends.

Strict application of copyright law prevents this degree of cooperation and innovation. This is because the useful lifespan of software code is relatively short, often no more than a year or two. The limited duration of copyright and patent does nothing to reduce this stifling effect upon innovation. Open Source licences make it possible for creators of programming code voluntarily to open up their work to the possibility of such collaborative effort. Of course, many Open Source products are neglected or weak, and only the biggest projects attract the huge armies of collaborators that Mozilla, Apache and Linux can boast. This includes paid labour from employees of companies who have a vested interest of one kind or another in seeing the Open Source tools in question succeed. Companies that want to keep their secrets close and their labour in-house simply do not use Open Source licensing; indeed, in some circles Open Source is almost a dirty word.

⁸ Text of GPL from <<http://www.gnu.org/copyleft/gpl.html>>; cf. the legal code of the Creative Commons' Attribution-ShareAlike licence, at <<http://creativecommons.org/licenses/by-sa/3.0/legalcode>>.

We would argue that like software, academic research cannot afford to wait for copyright to expire.⁹ In disciplines such as medicine and related sciences research more than five or ten years old is often considered, if not obsolete, then old news; in most cases it will have been tested, built upon and either superseded or rejected. Publications in the humanities do not have such short shelf-lives; it is not unknown for a work a hundred years old still to be cited in current research in the same area. Nevertheless, it is probably true that most engagement with a new piece of research, be it in a critical review, a refutation or a re-use of a theoretical model with different parameters, takes place within the first decade or so of its publication. It is essential that no legal barriers stand in the way of such free citation of and engagement with ideas.

It is recognized that critical editions are not new creations in the sense that copyright is meant to protect, but that they are, by definition, reflections and derivations of existing (usually public domain) material. Even in traditional publication media, both in printed editions such as Oxford Classical Texts and searchable corpora like the *Thesaurus Linguae Graecae*, it seems to be the assumption that verbatim reproduction of previous publications is, within certain limits, fair game. The apparatus criticus is another story: this is scholarly creation and is considered to be protected by copyright.

It is also of course clear that copyright does not prevent the ideas in a piece of scholarly literature from being cited, built upon and argued with, but only protects the actual words used in the expression of said ideas from being copied, redistributed and profited from by anyone but the author (or, more often, the publisher). In the digital age, however, there is more to scholarship than simply abstract ideas expressed in elegant rhetoric and language; sometimes the most essential part of an academic work is precisely the actual words and codes used in the expression of that work. This is true, for example, of critical editions, eclectic or supplemented texts, apparatus critici, philological commentary and text encoding, as we shall discuss below. A database or XML-encoded text is not merely an abstract idea, it is itself both the scholarly expression of research and the raw data upon which that research is based, and which must form the basis of any derivative research that attempts to reproduce or refute its conclusions.

In the case of digital editions, therefore, conventional copyright is arguably doing more to hold back research than to protect the author. We propose that a protocol for collections of digital critical editions of texts and/or manuscripts, which aims to allow for collaboration on the widest possible scale, must include the requirement (or at least the very strong recommendation) that texts are not only Open Content (allowing free access to the output itself) but also Open Source – revealing transparently the code behind the output, the research behind the text, the decisions which are part of scholarly publication.

⁹ See, for example, the discussion arising from and surrounding four articles on the value of Creative Commons licensing posted at <<http://creativecommons.org/weblog/entry/7435>>.

If a project were to publish digital critical editions without making the source code available, this would arguably be in conflict with the principles of scholarly editing and publication upon which the academy is based. Open Source in this context is not innovative, it is traditional.

Critical

‘Critical’, in this context, is a qualifier that is ostensibly clear in its denotation but betrays a more complex history and set of assumptions that will need to be elucidated, albeit briefly and selectively. Historically, ‘critical’ discourse in the sense we use it refers to nineteenth-century humanist scholarship, with roots reaching back to the Enlightenment and the Renaissance, to the Hellenistic librarians and earlier. In this tradition, criticism meant methodical assessment of evidence following well-founded criteria. Kant defined criticism not only as his contemporaries did, as a method of logical analysis, but more particularly as the absolute exercise of reason. In his *Critique of Pure Reason*, he called his own time the ‘Zeitalter der Kritik, der sich alles unterwerfen muß’.¹⁰ His own philosophy was aimed at establishing a firm foundation for knowledge, practice and emotion, and separating faith from knowledge.¹¹

Kant’s philosophy linked up with the establishment of the early-modern university, which gave criticism an institutional home. When, in 1793, the University of Halle was founded, an institutional model had been introduced that would soon spread throughout Europe and eventually, with considerable variation, all over the world. Unlike the orthodox universities in Leipzig, Wittenberg and Jena, which were controlled by the trans-territorial Lutheran church, the administrative structure of the University of Halle was reorganized by the Hohenzollern dynasty to make it accountable to the state ministry; to educate jurists, civil administrators, teachers and pastors primarily committed to the needs of civil society in the German state. As Ian Hunter puts it, the new institutional model was there to ‘divorce politics from theology and to fashion a style of thought and conduct that would allow jurists and administrators to subordinate the uncompromising ideals of religion to the peace and prosperity of the state’.¹² Hence, ‘critical’ implies the adherence to reasonable methods and principles, the radical judgement of interpretative decisions in accordance with those methods and principles and, ideally, the civil exchange of reasonable arguments between diverging opinions. It was the role of philosophy, according to Kant, to protect the university from undue influence from outside sources, be it from the state or the church.

10 I. Kant, *Kritik der reinen Vernunft* ([1781]; Hamburg: Meiner Verlag, 1998), p. 7: ‘the age of criticism, to which everything must conform’.

11 B. Recki, ‘Kritik’, *Religion in Geschichte und Gegenwart*, 4 (2001): 1781–2, p. 1781.

12 I. Hunter, ‘The Regimen of Reason: Kant’s Defence of the Philosophy Faculty’, *Oxford Literary Review*, 17 (1995): 51–85, p. 56.

When the University of Berlin was founded 17 years later, the modernization of the university had already made its decisive breakthrough and the German idealists, from Schiller to Humboldt, had entrenched the ties between the university and the nation-state, epitomized in the idea of ‘culture’. The study of culture, of course, required the interpretation of foundational texts, and Greek and Latin texts had been part of German *Bildung* for some time. In the seventeenth-century ‘critical’ editions had begun not only to be critical in their application of philological judgement, following the traditions of Alexandrian scholarship, but also increasingly critical in relation to the historical investigation of the sources and their relationships. It was, however, only around the person of Karl Lachmann that a set of rigorous rules was formulated in order to establish the oldest possible text from a group of manuscript witnesses. The ‘Lachmann method’¹³ set out the careful comparison (collation) of all extant manuscripts of a given work and the meticulous application of a series of steps that would lead not only to the constitution of an archetypal text, from which all manuscripts descended, but also the reconstruction of a genealogical family tree, which would delineate the interrelationship of all manuscripts deriving from said archetype.

Sebastiano Timpanaro described the four key traits of Lachmann’s *recensio* in which he had been influenced by a number of different Classical philologists:

[1] The rejection of the vulgate and the requirement that the manuscripts ... be used as the foundation of the edition. ... [2] The distrust for manuscripts of the Humanist period. ... [3] The reconstruction of the history of the text and particularly of the genealogical relations that link the extant manuscripts. ... [4] The formulation of criteria permitting a mechanical determination (without recourse to *iudicium*) of which reading goes back to the archetype.¹⁴

That an approach based on the rigorous application of rational methods could indeed reach conclusions unperturbed by, or at any rate in tension with, irrational forces outside the text can be illustrated by one of its most remarkable achievements: New Testament textual criticism initiated¹⁵ a project, still ongoing, to establish a ‘critical’ text of the New Testament, that broke with the entrenched

13 S. Timpanaro (*The Genesis of Lachmann’s Method*, ed. and trans. G.W. Most (Chicago: University of Chicago Press, 2005) and others have convincingly made the case that Karl Lachmann was probably neither the inventor nor the most consistent implementer of said method; we still retain this name out of convenience. See P. Maas, *Textual Criticism*, trans. B. Flower (Oxford: Clarendon Press, 1958) for a classic synthetic presentation of the Lachmann method.

14 Timpanaro, *Genesis*, p. 115f.

15 There were, of course, predecessors. Notably in the work of Erasmus (1516), Bengel (1734), Wettstein (1751/2) and Griesbach (1775/7), none of which fully succeeded in overcoming the text of the *Textus Receptus*.

Textus Receptus by strictly following the scientific rules set out by, amongst others, Karl Lachmann himself.¹⁶

At this point, it will be helpful to point out some central consequences of textual criticism – what makes a particular discourse ‘critical’ – pertinent to the argument put forward here. The critical method was drilled into generations of future scholars in seminars and closely supervised papers and theses. Later on, a system of reviews and discussions – ‘peer review’ – made sure that adherence to these rules was certain. The cumulative characteristic of this discourse, however, was that interaction with fellow critics – contemporary and past – had been put on the more civil basis of the exchange of rational arguments. There was no space, at least in the ideal case, for *ex cathedra* pronouncement or recourse to tradition, authority or dogmatics. Fellow critics had to be persuaded and they, furthermore, had to be persuaded solely by logical argument. Two further traits followed from this.

A crucial characteristic of critical discourse was – for obvious reasons, especially in the formative years – documentation of sources and preceding works. Many manuscripts were ‘discovered’ and edited by critical scholarship in the eighteenth and nineteenth centuries. Subsequent scholarship could base itself on those editions and improve them as further manuscripts were made available and improved readings of individual manuscripts and collations of a growing number of them were produced. Collecting – ideally all – preceding scholarship on particular subject matter in a bibliography and reviewing it critically became *de rigueur* in scholarly practice. One of the reasons for this was to identify clearly the contribution of an individual critical work vis-à-vis a growing body of scholarly discussion and achievement – the acknowledgement of the giants of scholarship on whose shoulders one stood. Another aspect of the relationship with antecedent work was to ensure a ‘fair use’ of other scholars’ publications. As mentioned above, scholarly output was, subsequent to its publication, open to use according to standards established and policed by the critical community.

Critical scholarship could lead to a variety of outcomes: a new thesis in answer to a particular question; but also a new or improved edition of a manuscript or text. These outcomes needed nonetheless to be persuasive to the scholarly community. Nothing detracts more from a convincing line of argument than a detailed critical interaction with the aforementioned body of scholarship, however necessary. This was the place for the apparatus criticus, a subordinate, though no less important, text alongside and closely inter-referenced with the main text, in which this interaction could be explicitly acknowledged and developed. The apparatus criticus could comprise any of the following: a collection of variant readings or translations, elaborate textual notes and discussion, adding to the already mentioned bibliography, index or even concordance. Nothing exemplifies this established apparatus better than the footnote. Leaving the main text elegantly clean and uncluttered by distracting references and argumentation, footnotes host

16 Lachmann’s own attempt, as Timpanaro relates it, was anything but a full success. It was for Constantine von Tischendorf to realize the first true breakthrough.

something which is fundamental to criticism. They are, as Anthony Grafton puts it, ‘the humanist’s rough equivalent of the scientist’s report: they offer the empirical support for stories told and arguments presented’.¹⁷ They provide the basis and infrastructure of critical discourse and thus disclose one important feature of critical practice and open up one crucial possibility. What this reveals is, of course, that criticism is fundamentally a *communal* enterprise. While the scholar’s argument is presented as the erudite position of, usually, an individual critic, what makes it possible in the first place is presented explicitly or implicitly in the apparatus. But in providing the basis of the research, including all references to sources used, it creates the possibility of verification to the reader, who is now potentially able to reconstruct, double-check and critically scrutinize each critical step that forms the basis on which the argument stands or collapses.

In view of the genealogical vignette and arguments above, one major implication of the qualifier ‘critical’ is, we hope, more than evident: text editions should only be seen as fully critical if all interpretative decisions that led to the text (on which more below) are made as fully accessible and transparent as possible. This is not to say that this ideal is always reached or that editions that do not adhere to this definition of ‘critical’ are to be disregarded. We should merely like to argue that full accessibility and transparency is the ideal toward which to strive if ‘critical’ is understood in the way laid out above.

The realization of the ideal itself, of course, has always been constrained by the medium in which the edition is presented. Practical consideration of costs, technical possibilities, possible layout and editorial tradition will always influence the degree to which all elements that forged the edition will be displayed in front of the reader. It is quite possible that all manuscripts will be fully transcribed, including records of all graphemic idiosyncrasies. Such palaeographic insights – often important for elucidating the text – have to be sacrificed, however, when grouping variant readings in the apparatus variorum. This is in itself a valuable critical act, but the critical steps leading to the groupings of variants, after they have been heavily normalized, are, in this case, no longer available to the reader who would like to reconstruct those steps. We should like to argue that a digital edition is less constrained than paper and print, in that it is possible to be inclusive of all critical arguments. In addition, both graphemic *and* normalized transcriptions of all manuscripts that comprise the edition could and should be provided, at least as one possible scenario. This is certainly not the only example of the economy of the print medium limiting critical discourse: previous discussions are often implicitly presupposed or, at best, merely alluded to. A selective choice of arguments is made necessary, exactitude and certainty are not always spelled out. All material potentially disrupting the overall argument in a commentary is relegated to footnotes or endnotes.

¹⁷ A. Grafton, *The Footnote: A Curious History* (Cambridge, MA: Harvard University Press, 1997), p. viii.

This potential explicitness is desirable not only in the edition's traditional narrative medium; many – though not all – critical decisions can and should also be formalized into a machine-readable form. Making this information machine-actionable transforms a digital edition beyond a mere emulation of a print edition; it opens the edition up to further computer-assisted analysis of text features and their complex relationship, as well as making it possible to recognize complex argumentation patterns. Critical transparency, however, works in both directions: most digital editions are encoded in XML or held in a relational database. XML schemas and entity relation models, as well as transformation scripts and query algorithms, are integral parts of such a digital edition. These editions are much more than the data visible on the user interface. If they are to be 'critical' in the aforementioned sense, all these aspects may also need to be made transparent and accessible.

We argue, in sum, that 'critical', in the context of digital editions, indicates not only the presentation of texts that are the outcome of critical scholarship but, more importantly, requires transparency, to software as well as to human readers, as to the editorial interventions made and the sources, data and scholarship behind such decisions.

Editions

Critical editions – even if 'critical' is understood in the more specific sense laid out above – come in many different types and forms. In the effort to explain this variety, Martin-Dietrich Gleßgen and Lebsanft have recourse to two conflicting principles. On the one hand, there is the principle of the concrete and factual materiality of the extant documents that survive in libraries, archives and museums; on the other, is the principle of the ideal and abstract notion of a reconstructed archetype, or even *Urtext*.¹⁸ Different editions are shaped by the tension between these extreme principles and are accounted for by the relative influence each principle is exerting upon the particular edition. The space between the extremes is therefore populated by a range of edition types that negotiate, mediate and compromise between these two foundational approaches. This begins with the artefact itself, and moves on to a variety of surrogates (drawn, photographed, scanned), 'diplomatic', normalized, synoptic and 'best-text' editions, until the full-on historical critical edition, complete with apparatus variorum, is reached. In order to elucidate the editorial issues raised from each pole, it might be helpful to look at two of their closest representatives in their print media incarnations: the papyrologist and the literary textual scholar.

18 M.-D. Gleßgen and F. Lebsanft, 'Von alter und neuer Philologie: Neuer Streit über Prinzipien und Praxis der Textkritik', in M.-D. Gleßgen and F. Lebsanft (eds), *Alte und neue Philologie*, Beihefte zu editio 8 (Tübingen: Max Niemeyer Verlag, 1997), pp. 1–14.

The papyrologist represents an editorial practice influenced by the principle of the concrete and factual materiality of ancient manuscripts. Herbert Youtie devoted his famous 1962 Russel Lecture to the more ‘private’ aspect of the editorial activity of the papyrologist: ‘The typical product of the hours spent by the papyrologist with his papyri is the edition of one or more texts. So much is this the case that we are ready to call a scholar a papyrologist if he publishes a papyrus text now and again, no matter what else he may do.’¹⁹ The reason that this activity is often neither appreciated nor fully discussed is, according to Youtie, owing to the fact that adjacent disciplines regularly accept published papyrus editions as ‘fact’: ‘The general accounts can afford not to tell us what the papyrologist does, because what he does is used up in producing texts that are absorbed into literature or history. It is self-consuming labor and leaves little or no trace of itself in the editions.’²⁰ Papyrologists know better than to merely trust the presented readings as facts, as they ‘manufactured’ most of these facts during the process of transcribing. The process of transcribing and editing a papyrus is not only amongst the most demanding critical philological activities in textual scholarship, it also necessitates a particular individual and collective workflow, in order to ensure the quality of the edition.

The processes leading to an edition of a papyrus manuscript are, in all experience, doomed to fall short of full success, since they ‘call for insight, ingenuity, and imagination to a degree no one man could possibly possess’. A host of factors, ‘the physical state of the papyri, the nature of the handwriting on them, the “dead” languages represented by the writing’, conspire to subvert the success of this endeavour.²¹ Papyrologists are fully aware of this condition and have developed ways to build on their early readings, either by publishing subsequent improved editions, or in *Berichtigungslisten*, which are collections with critical presentations of corrections and supplements to such editions. While the improvement of a published edition is made via sequential publications, these publications are merely instantiations of a cyclically improved collaborative intellectual product which involves ‘the laborious production of a transcript, the discovery of error, and repeated revision until all error is eliminated’.²²

In order to open up editions to such collaborative critical scrutiny, papyrologists have developed and adopted a detailed system. This system originated at the 18th International Congress of Orientalists (Leiden, September 1932) – the Leiden Convention.²³ It allows readers of such transcriptions, via a variety of marks

19 H. Youtie, ‘The Papyrologist: Artificer of Fact’, *Greek, Roman & Byzantine Studies*, 4 (1963): 19–32, p. 22.

20 *Ibid.*, p. 21.

21 *Ibid.*, p. 23.

22 *Ibid.*, p. 29.

23 On the congress, see especially B.A. van Groningen, ‘Projet d’unification des systèmes de signes critiques’, *Chronique d’Égypte*, 7 (1932): 262–9; A.S. Hunt, ‘A Note on the Transliteration of Papyri’, *Chronique d’Égypte*, 7 (1932): 272–4.

on the text such as brackets and other sigla and diacritical signs, to reconstruct which letters have been added, corrected or expanded by the editor, as well as the layout of the text on the original support: ‘From all [these] signs a practised eye can visualize the actual shape of the original text, locate its holes and tears, and perhaps even imagine its folds.’²⁴ Turner is, however, quick to add some of the shortcomings of the ‘Leiden system’, one of which is that ‘it is not possible to be so confident that what is stated to be there is there’. This shortcoming led to an alternative approach – the diplomatic transcription – being established alongside the application of the ‘Leiden system’. Turner describes this approach as follows:

This transcription does not separate words, and follows exactly the layout of the original for spacings and interlinear additions, accents, critical marks, etc. It shows no letter as read which cannot be guaranteed. Ambiguous traces are described in the critical note. ... One purpose of making such a transcript is to force the transcriber to discriminate between what he sees and what he would like to see, to call his attention to the subjective factor in decipherment, and to make him devise verifications for his readings.²⁵

In order to prepare a critical print edition, the papyrologist usually has to decide beforehand – unless working with a particularly generous publisher – what kind of edition will be presented: how much of the text to normalize, add and correct, and how much of the decision-making process to make explicit.

While the papyrologist edits a text still bearing the traces of its physical appearance on the actual manuscript, textual scholars seek to construct an altogether more disembodied work. As Paul Maas formulated it in the very first pages of his classic textbook, ‘the business of textual criticism is to produce a text as close as possible to the original (*constitutio textus*)’.²⁶ But this formulation begs at least two questions: that of how close one can get to the original; and, more fundamentally, how safe it is to assume that there was such a thing as an ‘original’. Martin West, established Classical scholar and editor of the Teubner *Iliad* edition, admits this difficulty in his editorial work on Homer: ‘The Homeric poems, because of their oral background and the special nature of their early transmission, pose peculiar problems to the editor and textual critic. To begin with, there is the problem of defining what the text is exactly that they are aiming to establish.’²⁷ The problem with the Homeric poems is that they have ‘grown’ considerably over the time of their transmission. There was neither a definable Urtext, nor a single ‘Athenian’ text, as quoted by Hipparchus, Thucydides and Plato. Not even Alexandrian

²⁴ E. Turner, *Greek Papyri: An Introduction* (Oxford: Oxford University Press, 1980), p. 70.

²⁵ *Ibid.*, p. 71.

²⁶ Maas, *Textual Criticism*, p. 1.

²⁷ M.L. West, ‘The Textual Criticism and Editing of Homer’, in G.W. Most (ed.), *Editing Texts: Texte edieren*, Aporemata 2 (Göttingen: Vandenhoeck & Ruprecht, 1998), p. 94.

scholarship, which probably established some agreement about its *Versbestand*, determined a single circumspect text. In view of this circumstance, West proposes a compromise: ‘Let us state our aim to be the establishment, so far as our means allow, of the pristine text of the poems in the form they attained following the last phase of creative effort.’ In other words, we must concede that, as the tradition has passed through several centuries of ‘wildness’, it may be impossible to establish exactly what lies on the far side.

In the reconstructions of the New Testament, on the other hand, proximity to the autographs is an ideological demand, since it is traditionally seen as the authoritative text written by divinely inspired authors. Any deviation from the original necessarily represents a corruption of the canonical text. In 1882, editors Brook Westcott and Fenton Hort could still give their critical edition of the New Testament the ambitious title *The New Testament in the Original Greek*. After the collation of numerous additional manuscripts and copious scholarly discussions, the Institute for New Testament Textual Research in Münster, responsible for the standard critical edition, has adopted the more realistic aim of reconstructing the *Ausgangstext* (‘Initial text’), i.e. ‘the text that precedes [the] process of copying’, adding: ‘Between the text of the author and the initial text there may be developments that have left no traces in any of the surviving manuscripts.’²⁸

Critical editions of texts with multiple manuscript witnesses have tended to present the reconstructed text as the lemma and a summary of agreements and disagreements with the individual manuscripts or manuscript groups in the apparatus variorum. But even though the aim of reconstructing the oldest possible text has been widely accepted as the core critical task, how to present it has remained an incompletely resolved issue. The classic representative criticizing the presentation of an eclectic text as lemma remains Joseph Bédier.²⁹ His disagreement with the Lachmann method demonstrated a profound uneasiness about offering a reconstructed text that is not attested in its wording in any concrete manuscript. His suggestion was to choose the text of a concrete manuscript that is deemed by the editor to most closely represent the archetype of the text tradition – the ‘Best Text’. ‘Best’ was still defined by Bédier as coming closest to the notional source text from which all other witnesses depart. However, what if a textual tradition is more complex than a tree model could explain? What if the transmission of the text implied not only copying but continuous creative rewriting leading to chronic instability within the text – variance. Bernard Cerquiglini in his 1989 essay *Eloge de la variante*³⁰ is chief instigator of a

28 Institute for New Testament Textual Research in Münster, available at <<http://www.uni-muenster.de/INTF/>>.

29 Famously discussed by Bédier in ‘La tradition manuscrite de Lai de l’ombre’, *Romania*, 54 (1928).

30 B. Cerquiglini, *In Praise of the Variant: A Critical History of Philology*, trans. B. Wing (Baltimore: The Johns Hopkins University Press, 1999).

re-orientation of the philological perspective. Basically, the ‘new philology’ can be seen as embracing the concrete and factual materiality of texts.

Digital editions may stimulate our critical engagement with such crucial textual debate. They may push the classic definition of the ‘edition’ by not only offering a presentational publication layer but also by allowing access to the underlying encoding of the repository or database beneath. Indeed, an editor need not make any authoritative decisions that supersede all alternative readings if all possibilities can be unambiguously reconstructed from the base manuscript data, although most would in practice probably want to privilege their favoured readings in some way.³¹ The critical edition, with sources fully incorporated, would potentially provide an interactive resource that assists the user in creating virtual research environments. Responses to a richer variety of analytical perspectives would be made possible and this would feed into future editions. This model should enable a more holistic notion of what is understood by a text as well as which sources can be represented by a modern edition.

Concluding remarks

To summarize, the model we are proposing here is for digital critical editions to be recognized as a deeper, richer and potentially very different kind of publication from printed editions of texts, even if such editions are digitized and made available in open content form. Open Source Critical Editions are more than merely presentations of finished work; they involve an essential distribution of the raw data, the scholarly tradition, the decision-making process, and the tools and applications that were used in reaching these conclusions. The Open Content model is an extremely important new movement in publication; the OSCE proposal is for a potentially new approach to research itself.

In theory the editorial and even publication implications of the Open Source Critical Editions discussion allow for a wide range of approaches, from a traditional one-editor text published in static form to a free-for-all wiki that can be contributed to concurrently and without restriction by any number of editors.³² However, as we have stressed above, our model calls very clearly for all editorial contributions, modifications and decisions to be transparent and explicit, to be attributed and citable, and to be stable and permanent. We do not have space in this chapter to

31 An example of such a radically ‘agnostic’ editorial policy is that of the *Online Critical Pseudepigrapha*, available at <<http://www.purl.org/net/ocp>>; all witnesses (including scholarly emendations) are transcribed, and eclectic editions can be generated from combinations of these at the presentational stage.

32 See, for example, Peter Robinson’s abstract titled ‘A New Paradigm for Electronic Scholarly Editions’ (2006) <<http://www.methodsnetwork.ac.uk/activities/es03abstracts.html>>, which seems to imply a relatively unstructured approach to scholarly editing of digital editions.

discuss all the protocol and workflow issues that editors of this sort of edition need to address, but we should like to stress that openness and critical method require both robust citation and versioning protocols, and for whatever editorial control is in place to be documented and fully attributed.

Clearly there are issues of technology, protocol and academic workflow that arise from the model proposed in this chapter. All of these issues are being explored and will continue to be explored both within the digital Classicist community and the much wider world that is digital academia. In particular there are two sets of issues that interest us, and which we feel warrant further exploration and discussion. The first of these is the need for the development of technologies, protocols and methodologies to make the OSCE model possible. Many of these may be adapted from those used in other disciplines. The second issue is the relationship between published OSCEs and large collections of less deeply marked-up but Open Content texts. It will be essential to ensure not only that the larger collections have the capacity to include the richer texts, but also that the protocols adopted by both movements allow the wealth of the smaller, deeply encoded body of editions to be used to enrich the collection as a whole and educate the technologies that query, organize and deliver it.

We should also stress that, as Crane has been arguing for years,³³ it should not be Classicists, or even perhaps digital humanists, who are inventing completely new technologies and protocols and workflow methodologies as our intellectual and academic world evolves. Many of these issues will have been addressed by other disciplines, in particular (but not only) the sciences, and where possible we should adopt or (if they are Open Source, for example) adapt the tools of these better-funded fields. There are gaps in the resources available, and some of our needs are either unique or as yet unmet, but these gaps are better filled by adapting from and contributing to the wider academic community than by inventing new methods from scratch, just as the EpiDoc Guidelines for publication and interchange of Greek and Latin epigraphic documents in XML built upon the solid groundwork provided by the venerable Text Encoding Initiative rather than creating their own schema from nothing.³⁴ It will not always be possible to borrow wholesale from other disciplines, but Classicists should take advantage of models, tools and systems developed in other fields, without underestimating the extent to which the approaches designed or improved within their own framework might be fed back successfully to the lively exchange of cross-disciplinary ideas and experiences.³⁵

33 G. Crane, 'Classics and the Computer: An End of the History', in S. Schreiban, R. Siemens and J. Unsworth (eds), *A Companion to Digital Humanities* (Oxford: Blackwell Publishing Limited, 2004), pp. 46–55.

34 All EpiDoc materials, including guidelines and documentation, accessible via <<http://epidoc.sourceforge.net/>>; history and current publications of the Text Encoding Initiative at <<http://www.tei-c.org/>>.

35 The Canonical Text Services (originally 'Classical') is an excellent example of such a technology created specifically by and for the Classical field, but which draws from

We hope that presenting the rationale of the OSCEs by way of unpacking three of its core concepts will provide a summary of a vigorous discussion of these issues, open these arguments up to a wider audience, and lay the foundation for future projects and discussions.

information science principles (e.g. FRBR), and should be taken up by scholars with a concern for text repositories from all backgrounds; see Christopher Blackwell and Neel Smith, 'A guide to version 1.1 of the Canonical Text Services Protocol', available at <<http://chs75.harvard.edu/projects/diginc/techpub/cts-overview>>.