Open Library of Humanities

Part of the Ubiquity
Partner Network

Peer Review:

Copyright:

Open Access:

Digital Preservation:

(dm/)

REVIEW ARTICLE

On Not Writing a Review about Mirador: Mirador, IIIF, and the Epistemological Gains of Distributed Digital Scholarly Resources

Joris van Zundert

Huygens Institute for the History of the Netherlands Royal Netherlands Academy of Arts and Sciences, NL

joris.van.zundert@huygens.knaw.nl

This piece mushroomed from a simple enough looking suggestion to write a review about Mirador, a viewer component for web based image resources. While playing around and testing Mirador however, a lot of questions started to emerge—questions that in a scholarly sense were more significant than just the functional requirements of textual scholars and researchers of medieval sources for an image viewer. These questions are forced upon us because of the way Mirador is built, and by the assumptions it thereby makes—or that its developers make—about its role and about the larger infrastructure for scholarly resources that it is supposed to be a part of. This again led to a number of epistemological issues in the realm of digital textual scholarship. And so, what was intended as a simple review resulted in a long read about Mirador, about its technological context, and about digital scholarly editions as distributed resources. The first part of my story gives a straightforward review-like overview of Mirador. I then delve into the reasons that I think exist for the architectural nature of the majority of current digital scholarly editions, which are still mostly monolithic data silos. This in turn leads to some epistemological questions about digital scholarly editions. Subsequently I return to Mirador to investigate whether its architectural assumptions provide an answer to these epistemological issues. To estimate whether the epistemological "promise" that Mirador's architecture holds may be easily attained, I gauge what (technical) effort is associated with building a digital edition that actually utilizes Mirador. Integrating Mirador also implies adopting the emerging standard IIIF (international image interoperability framework); a discussion of this "standard-to-be" is therefore in order. Finally the article considers the prospects of aligning the IIIF and TEI "standards" to further the creation of distributed digital scholarly editions.

Keywords: Mirador; IIIF; distributed knowledge; digital scholarly resources; information architecture; epistemology

Mirador at First Glance

Mirador (http://projectmirador.org/) is an open source, web-based, general-purpose image viewer written in JavaScript. Rashmi Singhal of Harvard Arts & Humanities Research Computing (http://darthcrimson.org/people/, https://github.com/rsinghal on Github) and Drew Winget of Stanford University Libraries (https://medium.com/@aeschylus, https://github.com/aeschylus on Github) are the chief authors of the code, although some forty-five additional developers contributed to the codebase according to Github (https://github.com/ProjectMirador/mirador/graphs/contributors). The development of Mirador was made possible by a grant from the Andrew W. Mellon Foundation to Stanford University. Creating Mirador has been—and still is—very much an open source and community effort, that also involves, for instance, well known scholarly open access advocates such as Robert Sanderson of the Getty Foundation and Loyola University of Maryland based Jeffrey Witt, known amongst other things for the Scholastic Commentaries and Texts Archive (http://scta.info/).

So, apart from an impressive open source community-driven project, what does Mirador deliver? In what follows I examine and use Mirador 2.1.2. It should be noted that, at the time of reviewing, a version 2.4 had just been released, but this version was not yet compatible with some of the additional tools that I wanted to experiment with in connection to Mirador. Meanwhile versions 2.5 and 2.6 have been released, and according to the change logs these versions offer better compatibility and backend support for (remote) annotations, an important aspect I wanted to examine. I have not re-examined this aspect with newer versions. The reader should thus take care to compare statements about Mirador with current developments, especially as Mirador is meanwhile preparing for a new life cycle and version 3.0 is expected to be released end of 2018.

If Mirador has been integrated in a website with a collection of images it actually provides a lot of image viewing capabilities right out of the box. I will discuss Mirador primarily from the perspective of a textual scholar interested in transcribing historical manuscripts, but the viewer is certainly not limited to showing facsimiles. Because Mirador is agnostic to the kind of images one uses, it can basically be tailored to the

needs of any scholar wanting to study images of objects—historians, art historians, book historians, musicologists, literary researchers, and so forth—so long as two-dimensional images will do. Mirador lacks any 3D viewing and rendering capabilities, so it caters best to people who want to study and compare paintings, facsimiles of manuscripts, images of typeset pages, photographs, music score, basically anything that was intended to lead its life in a 2D medium.

Mirador is very configurable, and the initial view will therefore vary. A typical setup, however, might not be unlike what is depicted in **Figure 1**. Out of the box Mirador offers convenient panning and zooming, and some controls to adjust basic image aspects, such as rotation, contrast, and saturation (cf. **Figure 2**).

For the purposes of showing Mirador's capabilities I have opted to use a few folios from a digitized medieval Dutch manuscript, the so called Comburg Manuscript (Comburger Handschrift, Württembergische Landesbibliothek Stuttgart, shelfmark Cod.poet.et phil.fol.22). It contains a collection of some well-known Middle Dutch works. The folios shown contain the first columns of the Middle Dutch fable *Van den Vos Reynaerde* (transl. *Of Reynaert the Fox*), of which a full translation is also open access available in English (Bouwman and Besamusca 2009).

If a collection of images is being viewed, controls to leaf through the collection are also provided: left and right pagers, and a thumbnail bar. The latter can be collapsed to save screen estate for viewing the actual image (cf. **Figures 3** and **4**). In
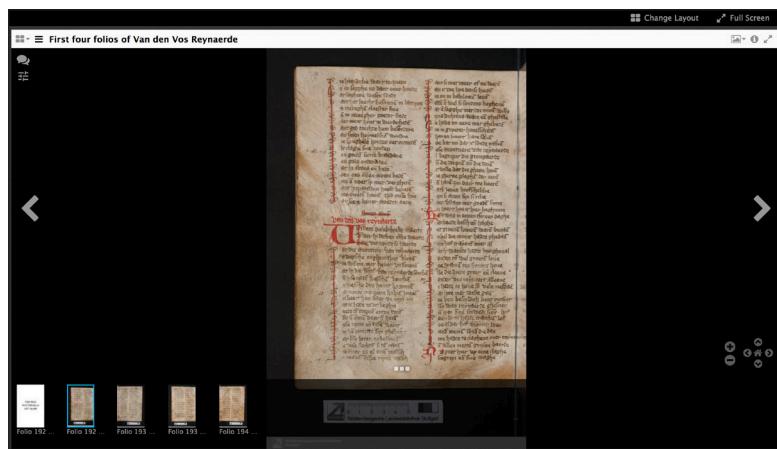

**Figure 1:** Mirador "Out of the Box".
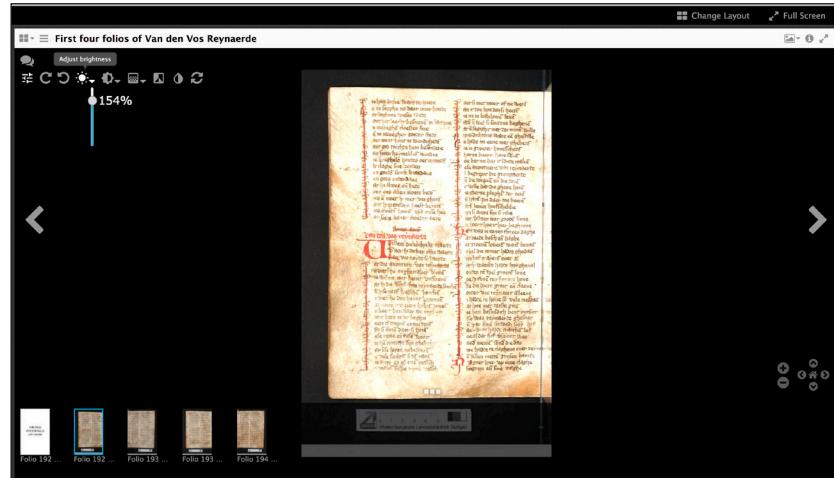
Figure 2: Item, with image controls expanded.
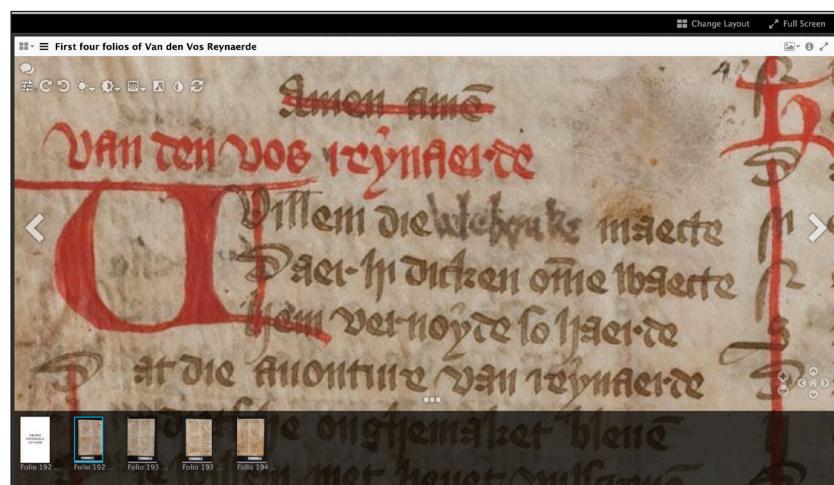
Figure 3: The same image, zoomed and panned. Image rendering controls expanded at the top left.

the case that a collection has properly added metadata, Mirador will also instantly display an index/contents, if so configured (**Figure 5**).

Mirador's technical designers and developers ought to be commended for not suffering from the "not invented here" syndrome. Instead of developing a completely new code base, they have maximally reused existing software components. This is a sensible development strategy as it lessens the development burden and prevents
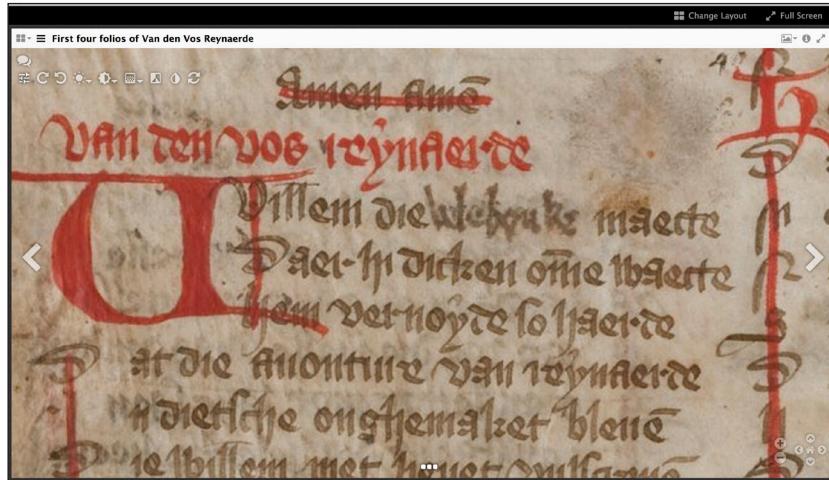
Figure 4: Same, with thumbnail strip collapsed.
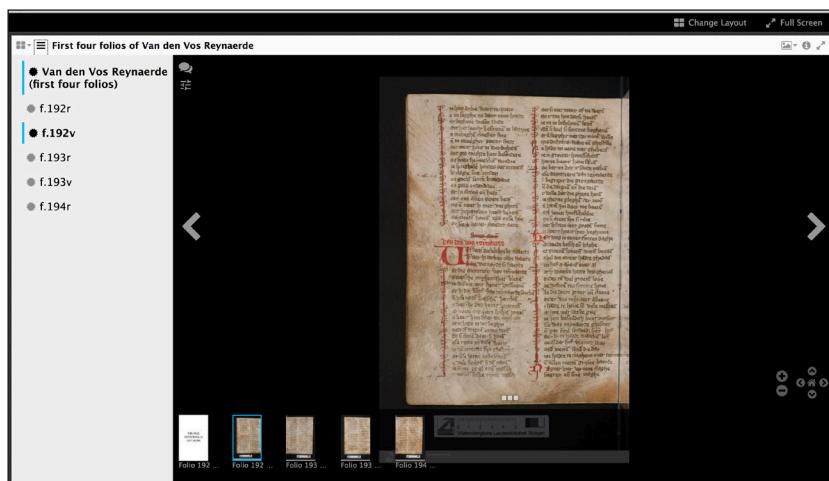
Figure 5: Mirador viewer with index visible.

the reproduction of many bugs, technical pitfalls, and maintenance issues. It is also a development strategy that is in line with the nature of open source and community-based software development.

Essentially then, Mirador wraps a number of pre-existing software libraries together and tries to turn that combination into a general-purpose image viewer. I will defer judgement for now on whether that attempt was successful. First let us see what is in the box. Mirador encapsulates the OpenSeadragon (https://openseadragon.

github.io/) viewer which delivers image viewing, zooming, and panning abilities. OpenSeadragon ensures a seamless and high-quality viewing experience. Under the hood JQuery (providing GUI controls plus a score of general purpose code) and TinyMCE (a powerful "rich text" editor) are code libraries that by now can be called part of the furniture in web development. A few other utility libraries are also encapsulated by Mirador. One reused component with a more central role is the Isfahan.js window manager (https://github.com/aeschylus/Isfahan), which itself uses a tiny part of the well-known and currently very popular d3.js (https://github.com/d3/d3) data visualization library that powers many a flashy dendrogram, wobbly network graph, and shiny bar chart. This window manager is a quintessential cog of Mirador. A window manager is the software component that allows you to open, resize, and move around windows on your computer's screen. Isfahan does the same for windows inside your browser. Mirador reuses Isfahan to allow the user to open and arrange an arbitrary amount of image viewers within the same browser window. This drives one of Mirador's self-proclaimed paramount features: the ability to compare images. Given that comparing codices and manuscripts is the bread and butter of textual scholarship, there is utility in being able to put two, three, or more codices next to each other on one's screen. So, if a Mirador viewer instance is set up for a particular collection of manuscripts, additional "slots" can be added for viewing other folios in the collection (**Figure 6**).

Of course, adding more and more views of manuscripts will soon turn a workbench into an impressively cluttered graphical interface, even if the user happens to have a 5120 by 2880 pixel monitor with 27″ screen diagonal—which is unlikely as roughly half of that (1322 by 768 pixels on a 15″ screen) is much more middle of the road currently. As Mirador is highly configurable the user may want to remove all that clutter of index buttons, paging buttons, thumbnails, and image property controls in order to attain, as the documentation has it (http://projectmirador.org/docs/docs/getting-started.html), "Zen mode" (**Figure 7**).

From a scholarly point of view, all that viewing "power" at one's finger tips is a dream. But I am a pretty spoiled scholar/developer, and therefore I still miss a *locked scroll* or *parallel pan* feature which would scroll or pan both (or more)
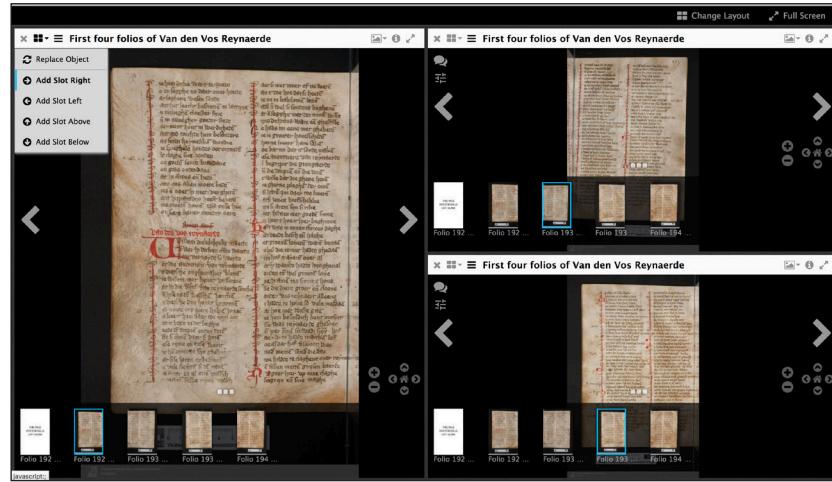
Figure 6: Multiple folios open for comparison.
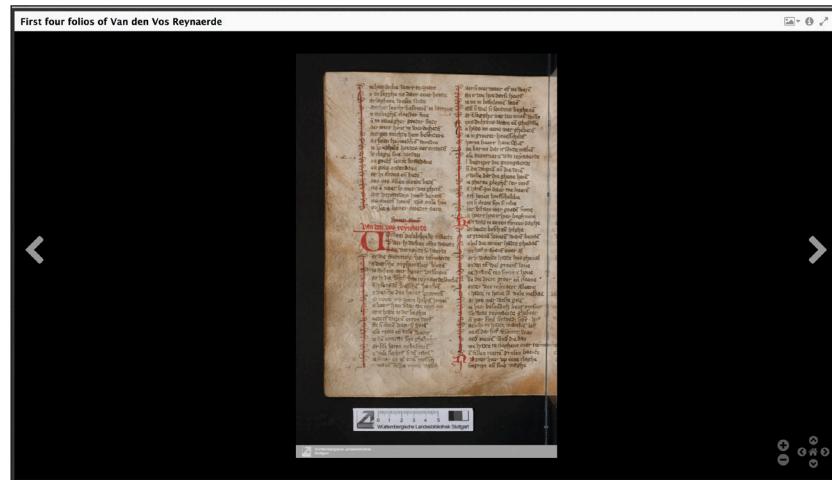
Figure 7: "Zen mode".

images that I am looking at simultaneously. This would be a feature that makes comparing lines a far less tedious task. (It should be noted that although this feature is not available out of the box there is a plugin that exactly provides this behavior.)

Another arguably indispensable feature of Mirador is the ability to annotate images, for which a convenient tool ships with Mirador by default. A scholar can draw a border around any arbitrary area and add annotation text for that area (**Figure 8**).
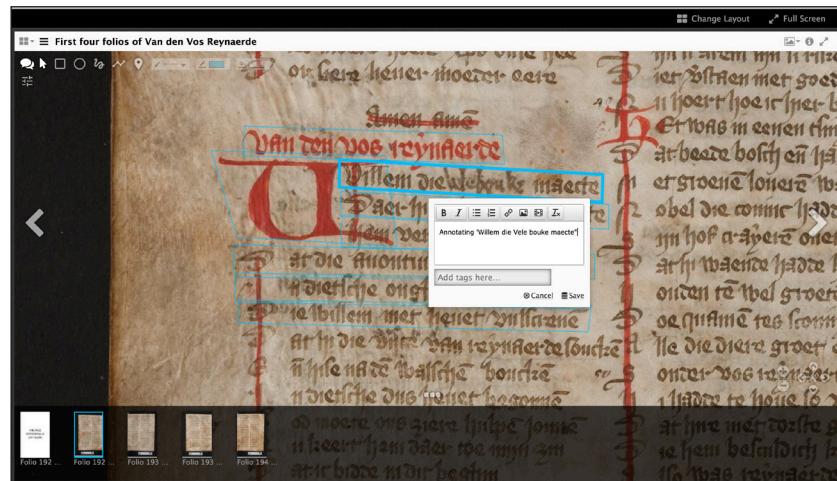
Figure 8: Annotation creation in Mirador.

In all, then, Mirador delivers an impressively well-functioning and rich viewer and annotation tool. Panning and zooming is all rather nicely instant and seamless, which makes for a comfortable viewing experience.

How not to write a review for Mirador

Of course, I could have called it a day right there. This review was to be handed in by January 2017. Theoretically I could have made that deadline, I think. But the text above still felt improperly incomplete as a review of Mirador. Nothing in there is really factually wrong—but it is also a far cry from the larger story of which Mirador is a chapter. Also, that larger story needs to be told. But "How hard can it be?" is a dangerous thought, and I never saw the end of it. Let me tell you about it…

Given the wrapping and encapsulation of ready-made components that Mirador does, one could critique it for being "just" a thin graphical user interface over a set of pre-baked code libraries, adding little of essence that could not have been achieved by several other means. However, first of all such a judgment would not be fair to the effort it must have taken to integrate all these libraries and functionalities. But more importantly: it would not acknowledge the pivotal role Mirador plays in what may be no less than a paradigmatic shift in how we understand, approach and interact with cultural heritage resources. Clarifying this will take a bit of explanation.

The Persistence of Silos

The ultimate transcription environment has become something of a holy grail within digital textual scholarship. Many attempts have been and are being made to create a transcription environment that surpasses any other environment, up to and including text editors, including MS Word (the surpassing of which arguably should not be too hard an achievement). Plenty of integrated transcription environments therefore exist. EPT (Kiernan, Dekhtyar, Jaromzcyk, and Porter 2004), T-PEN (http://www.t-pen. org), TextGrid (https://textgrid.de/), eLaborate (http://elaborate.huygens.knaw.nl/), and CTE (http://cte.oeaw.ac.at/) are some of the ones that I know to exist or to have existed, and I know of several others in a "pre-beta" state of perpetual development.

There is also a graveyard somewhere for scholarly transcription environments. Failed projects almost never get proper epitaphs or eulogies—the one for Project Bamboo by Quinn Dombrowski (2014) being the notable exception—which is a pity because such eulogies would be highly informative. In the case of transcription tools, a defining trope would be that the tool was built as an "integrated transcription environment." Integrated is a more formal term for "does it all." Their developers often want these tools to be the start and end all of digital textual scholarship work. This most often means that these tools expect all resources to reside in one place— specifically, on the computer or server where the tool is deployed. The reasons for this are not technical, born instead of institutional necessity and development convenience. The institutional makeup of academia and its (grant) funding schemes favour local institution-level digitization and development (Prescott 2016). Collaborative development between institutions is often frustrated by funding limitations, and moreover requires significantly more coordination effort than local development. Lastly, from the point of view of the developer, it is simply convenient to have all data in the same form and format and at an arm's length. Just as it is convenient for a scholar to have all sources and secondary literature on her desk, this saves a lot of tedious logistics needed to gather and process information. The effect of convenience for institution and for developer however is that these tools turn into so-called "data silos" (https://en.wikipedia.org/wiki/Information_silo): all

images, texts, annotations, and so forth need to reside on the same server to be used by the tool.

This has been a long-standing problem in digital scholarship technology. As far back as 2007, when a group of colleagues and I began a project called Interedition (http://www.interedition.eu/), the situation was almost perfectly the same: almost every institution, in some cases even each individual professor in textual scholarship was somehow involved in creating a large integrated all-purpose research environment. This caused (and causes) a lot of reinvention of wheels and duplication of effort, in a field that is notoriously understaffed with digitally and computationally skilled scholars and developers. At the time, it seemed to us—a mixed group of digital humanities developers, researchers, and any hybrid form in between—a good idea to reuse tools and resources rather than having local copies of text files and annotations, local tools, and the local burden of integration and graphical user interface development. This was not just a developer's concern, to keep development loads small; it also seemed to us that keeping tools and data locked in one place behind a one-size-fits-all interface would be in stark contrast to the heterogeneous demands and requirements scholars had for document resources that were spread, quite literally, across the planet. What use would it be to have an alignment tool locally in Würzburg if one of the documents it needed to align was in the Bibliothèque Nationale in Paris? Our buzzword of the day became "interoperability." The ideal was that no matter where you were, the tool at your disposal would work just as easily on a digital manuscript facsimile in, for instance, Florence as on a print edition in New York. We reasoned along the lines of service-oriented architecture. That is, distributed resources would be reachable via the same technical protocol language, which would guarantee that any local interface speaking that access language would be able to approach and use them. In that way it would not matter if I were using T-PEN and my colleague in Berlin were using EVT (https://visualizationtechnology.wordpress.com/), we would still both be able to hook into the same resource in Stanford. In 2011 my colleague Peter Boot and I also argued for such services-based digital scholarly editions in a more academic fashion (Boot and van Zundert 2011). It turned out, however, that between dream and reality stand institutional politics and practical considerations.

Practical techniques and methods for decentralization and resource reuse have been around for decades. The base protocol of the Web (i.e. HTTP https:// en.wikipedia.org/wiki/Hypertext_Transfer_Protocol) is intended to make both local and remote resources reachable from one and the same location—that is in fact the whole point of the Web (Berners-Lee 1995). Roy Fielding's 2000 definition of a Service Oriented Architectural style (SOA) to interoperate digital resources on the Web, called REST (https://en.wikipedia.org/wiki/Representational_state_transfer), did much to facilitate the creation of lightweight, small, and easily maintainable web services, based on Internet technology that had been around since 1994. REST made it easy to create Web APIs (Application Programming Interfaces), which in turn made it easy for developers to create web clients that could talk to more than just one specific server and vice versa. Suddenly, it seemed, there was a shared syntax and vocabulary to allow web applications to interact. It was a solution that was simple and obvious (at least to web application programmers). Unfortunately, although it is a necessary pre-requisite, a suitable technology is not in itself sufficient to change an institutionalized tradition of walled-in local digital resources and specific local methods of working with those resources. Thus, after the better part of two decades, data silos are still around, even if almost all scholars, scientists, and librarians agree that sharing data and documents in principle is a virtue (Fecher, Friesike, and Hebing 2015).

Why Are DSEs Still Information Silos?

Arguably, the majority of digital scholarly editions (DSEs) are still data silos. Browsing through the issues of RIDE (http://ride.i-d-e.de/), a reader sees exclusively digital editions that are fully locally-integrated server applications. The excellent Catalogue of Digital Editions (Franzini, Andorfer, and Zaytseva 2016 https://dig-ed-cat.acdh. oeaw.ac.at/) also lists for the most part API-less digital editions—12 out of 258, or less than 5%, do have an API—and whether those APIs are sufficient to "un-silo" the editions is unclear. Why has nothing changed in this situation in over two decades? As has been argued above, this is not a problem of technology. The institutional landscape and development convenience may be part of the explanation. However, sharing digital resources does not require collaboration per se. For a digital edition to share its resources with the world, it only requires unilateral action by its editor. The simplest thing that could possibly work is allowing web directory access to a server

location that hosts a dump of the digital materials—a technical no-brainer really. Such open data editions however, do not seem to thrive. Why?

I have at least two contentions about this. And they are truly contentious; I have only experience and hearsay to back them up—no real data, survey, or statistical analysis. But I offer these as the gambit of academic debate. My first contention is that textual scholars are still deeply entrenched in an intellectually hedonistic ideal of publishing the definitive edition. Most editors think of an edition as a complete and finished product—something that should not be tampered with, because it is argued and polished with arduous effort to academic perfection. The idea of reuse of that edition as "data"—especially as "primitive" or "raw" data in some computerized algorithmic process—in the eyes of scholarly editors is a category error. Because they do not regard that particular form of reuse as viable, textual scholars also do not expect textual resources to be offered as such. In other words: there is no innate wish within the community to push for a more distributed and interoperable model for text resource reuse. Even if it would be convenient for textual scholars themselves to be able to compare manuscript A in Rome with manuscript B in Zürich from their desk in London, that convenience is not a sufficient motivation to strive for some generalized method for decentralized access to textual resources, because that simply is not part of the teleological worldview of the textual scholar. Plus, I think, most textual scholars that have produced a digital scholarly edition would argue that it can actually be reused: look, it is a website, you can look at it, read it, use it. Again, however, that is a limited teleological conception of resource and reuse. It is the edition offered exclusively as a whole, as a solid philological fact, which is very much situated. A scholarly edition should be treated as a time-, context-, and editor-dependent collection of interpretations, rather than a set of "philological facts"—which, according to Jerome McGann (2015), are recorded, objectified, and archived inscriptions of documents, but which I doubt truly exist as fact. In contrast to this monolithic teleological view of the (digital) edition, colleague scholars seldom need the edition as a whole, but more often a part thereof. They want to compare a particular reading in one text with a different reading in another, or want to compare concepts and ideas over a range of documents, contrast different historical

perspectives on particular events described differently in subsections of different codices. With the digital scholarly editions we have, this requires scholars to navigate a plethora of different graphical user interfaces and differently-implemented search tools, differently-visualized annotation sets, etc.

My second contention is that the creators of digital editions—be they developers, scholars, hybrids, or any team-like combination thereof—veritably never have the resources (time, funding, or skill-wise) that enable them to produce the type of webservice-based edition that Peter Boot and I, as well as many others (Robinson 2013; Siemens, Crompton, Powell, and Arbuckle 2016; Thiruvathukal, Jones, and Shillingsburg 2010, and so forth) were thinking about. It consumes enough time and effort to create meticulous transcriptions of costly digitized manuscript folia in TEI-XML, and to put these together in some concerted form to bring them to the Web. If those priorities are met, there is usually very little capacity left to tend to luxuries such as machine access to the digital results.

Iron Manning the Digital Scholarly Edition

If the two contentions above hold water, they go a long way to explain why digital scholarly editions are all still based on a model of a unique, undivided, and complete product, and why Robinson's (2004) dream that "all readers may become editors too" by reusing the various parts of digital editions, and creating their own transcriptions and annotations, has not taken off. The second contention stands as a model for a great number of practical and pragmatic choices and limitations that influence the process of digital scholarly editing and the result of that process. Here however, I am not interested in these practical considerations. Nobody denies that it is highly convenient to have a text available anywhere, anytime. Nobody denies that it is practical to avoid travelling to the national library of another country to check particular artefacts on a specific folio. No one has denied the facilitative nature of a full text search. And every editor is aware of the limitations that funding, institutional policies, and capacity put on any edition. Yet still the obvious practical, pragmatic, and (possibly) financial benefits of publishing scholarly editions digitally have not convinced the majority of scholarly editors that digital editions are worthwhile.

I wonder how far this may be due to the particular stance that we, digital scholarly editors and computational humanists, have taken in arguing the digital scholarly edition. Advocates like myself have been hammering away at the practical advantages and revolutionary paradigm-shifting nature of electronic editions for more than a decade. Julianne Nyhan and Andrew Flinn (2016) argue that there is a distinct and overused revolutionary motif in digital humanities publications. Arguably this motif has worked as a shibboleth for the community, but it also may have worked rather counterproductively against the acceptance of digital methods in the humanities "proper", where the need for and advantages of a revolution were greeted with justifiable skepticism. Peter Robinson's "The Digital Revolution in Scholarly Editing" (2016) is an interesting publication in this respect. Clad in revolutionary terminology, it is highly critical of the aspects of digital scholarly editing usually depicted as revolutionary. Robinson argues that neither what we do as digital textual scholars, nor what we make constitutes a revolution in scholarly editing at all. We are still editors of texts after a critical fashion. Digital resources and environments may scale our work, but essentially do not warp us into any undiscovered paradigm. But Robinson continues to argue that there is still a truly revolutionary aspect to digital textual scholarship. It changes who we (textual scholars) are:

> Every edition I have discussed so far has been made according to what we might call the Alexandrian consensus. The librarians gathered the many texts of Homer together; the scholars studied them and created a text; a grateful world took that text and read it. This model rests on two pillars. The first pillar is that only qualified scholars may examine the primary documents. The second pillar is that only qualified scholars have the authority to make a text the rest of us may read. Both pillars are now fallen. We are moving to a world where every manuscript and every book from the past is online, free for anyone to look at. You no longer need to be tenured and well-connected to see a manuscript: increasingly, all you need is an internet connection. As for academic authority: peer-review and tenure committees are fine things but no-one is going to assert that only approved scholars can read manuscripts. (Robinson 2016, 198)

Robinson takes my second contention above (the perpetual lack of capacity, funding, and skilled personnel) as a major argument in favour of shared open digital editions. The facsimile materials should be put on the Web publicly under the freest license possible, open to all to transcribe, annotate, interpret, copy, perform, and so forth—an admirable altruistic and democratic argument, further underpinned by the fact that our work is usually financed through tax money.

The issue here is not whether I agree with Robinson, but his proclamation that this "revolution" is a fact, that the future course of (digital) textual scholarship cannot be but this one, and that it is already happening. Much academic literature about digital textual scholarship seems to subscribe to a similar premise. Franz Fischer's excellent contribution in Speculum (2017) postulates:

> the (albeit slowly) growing number of digital critical editions increases the demand for assembling and providing critical texts that are in the form of a textual corpus, because only collections or corpora of texts that are otherwise dispersed on various websites allow for a systematic analysis and for efficient research across the works of a specific author, genre, subject, period, or language as a whole. (S266)

Again, it is not whether Fischer is wrong or right here—in practice, I actually agree with his contention. What I find interesting is that Fischer postulates a world of existing digital editions (or data) and then suggests several solutions for how to reconcile the heterogeneity and specificity of critical editing with the homogeneity of digital corpora. These solutions are primarily methodological in conception, but with elements of IT architecture mixed in.

I want to cast Robinson's and Fischer's premises here as "iron manning", which is an inverse form of making a straw man argument (Casey 2012). This is unfair because their argument is well-intended and not necessarily wrong, but for the sake of argument I use their premises as examples of what advocates of digital humanities often seem to do: assert some digital ideal and make a case for what, methodologically, would fit this idealistic situation. Fischer and Robinson construct situations that are ideal for their reasoning, respectively that there are ever more

digital editions and that editions should be publicly open and sharable. These are subjective practical and pragmatic ideals to start with. One might wonder however, if practical and pragmatic ideals resulting firstly from a digital medium and only secondly from scholarship are appealing to conventional textual scholars. Are textual scholars not primarily interested in what can be known about a text, and should we not therefore demonstrate the added epistemological value of the technologies we propose first of all?

What is the Epistemological Case for Digital Scholarly Editions?

Until now, a paradigm shift in scholarly editing towards open and distributed digital editions seems not to have happened. Do readers really want to be editors? Moreover, most editors seem to be reluctant to embrace different, more open methods. The arguments cited in the previous section strike me as deterministic. That a digital technology exists and a futuristic ideal can be based on it does not necessarily mean that particular utopia can and must be reached. Could it possibly be that most textual scholars judge that the model of the singular unified complete text—or codex in any case, be it a digital or print one—is simply sufficient, or even superior to any open and machine-readable digital model yet presented? Fair is fair: we have little evidence that scholars and readers see a reason for a shared, interoperable, distributed model for text resources. At best there is evidence of the contrary in many a failed tool, the lack of use being made of digital editions (Porter 2013), and the persisting preference for print over digital by promotion and tenure committees (Schreibman, Mandell, and Olsen 2011). The open and social edition that is passionately argued for by Ray Siemens and others (Siemens, Crompton, Powell, and Arbuckle 2016) is a prototype at best and does not yet attract a mainstream audience. Proponents stress the practical aspects of open editions all the time, and especially the pragmatic democratic character, but they never explain what the epistemological gain is that justifies the extended trouble of creating open digital editions. Nor do they deeply consider what may be lost, as some encourage us to do (Sondheim, Rockwell, Ruecker, Ilovan, Frizzera, and Windsor 2016). Putting tax-financed editions digitally in the public domain is ethical, but that has no bearing on our epistemological grasp of the subject matter

or the historical objects that documents are. Proponents point to the "wisdom of the crowd", but that is also not revolutionary, either from a pragmatic or an epistemological point of view. It was already perfectly possible to write a "letter to the editor." Most editors—those who do not edit the works of Darwin or Dante's *Commedia*, but works more akin to some obscure and opaque 12th century book of prayers—would be happily surprised to receive one. The potential public reach of these works does not justify the development of heavyweight digital infrastructure or applications on the off chance that there may be an interested individual out there who actually knows more about an edition than the editor learned in the fifteen odd years spent studying it.

Calling it a revolution does not make a strong epistemological case. So, what could the epistemological case be? Why do we not discuss this widely in our field? Again, as a possible opening for a debate: two contentions. The first is that distributed open digital editions improve quality, and higher quality information advances knowledge. What constitutes quality of information is another conundrum that I will not detail here; suffice to say that it is also a highly situated concept (cf. Borgman 2015; Gitelman 2013). But the epistemological argument for distributed information lets itself be construed rather easily. It is connected to skill. Suppose I am a scholar in need of high-quality digital facsimiles of some folios of some medieval manuscript. I could try to obtain high-quality photographs and digitize them myself. But soon I would be facing questions like "How many DPI and what colour depth should these images be?", "How and where should they be stored?", "What is a feasible standard for the technical description of these photographs?". Chances are, a textual scholar is less equipped to answer these questions than a library-based digitization expert. The quality of the production of such digital facsimiles is related to the skill and knowledge of the producer. In contrast, if I want to be assured of the best possible transcription I am going to take my chances with the textual scholar specializing in 12th century European paleography, rather than with the librarian. This difference in assured quality of information does not evaporate post production. The curation and maintenance of digital information is yet another field of expertise, best left in the skilled hands of people maintaining

some digital repository. Thus, digital scholarly editions may be sites of intersecting knowledge that affirm and support specific and highly-skilled expertise. This does not mean they cannot be altruistic and democratic at the same time, opening up editions to the public, but the primary epistemological scholarly gain seems to be in better, more specific support for quality knowledge and expertise.

My second contention in support of distributed information is related to distributed knowledge, also known as group knowledge or, indeed, "wisdom of the crowd" (https://en.wikipedia.org/wiki/Wisdom_of_the_crowd). As pointed out above, arguing some epistemological advantage based on the "wisdom of the crowd" seems a dodgy fad at best, but it is a well-known fact that distributed knowledge adds up to more than local or individual knowledge (Fagin, Halpern, Moses, and Vardi 1995) and it can be explained relatively easily. Suppose some person X knows that fact A is related to fact B, but she does not know that fact B is related to some other fact C. Suppose also that another person Y knows that fact B is indeed related to C, but he in contrast does not know that B is related to A. The distributed knowledge that neither person has is that all three facts are related. They could gain this knowledge if the local information were somehow exposed and eventually shared (e.g. through word-of-mouth or publication). This kind of transformation of distributed knowledge into added knowledge is actually what textual scholars do all the time. What is, quite inexactly (Timpanaro 2005), called Lachmann's method is an excellent example of this. A scholar may find two copies of a text sharing the same copying error—in other words, information distributed over two sources. Combined, this information adds the knowledge that these copies are in all likelihood more closely genealogically related than copies that do not have that error.

The salient point here is that, although connecting distributed information might be done computationally, it must currently be done by hand, because the information that digital scholarly editions hold is represented almost exclusively through visual interfaces. This means that the epistemological benefit they can have depends on human agents connecting the dots. These human agents are part of a social epistemology (Goldman and Blanchard 2016), and distributed knowledge may certainly be uncovered in such a system of networked knowledge. However,

if the information within these silos were to be exposed in a way that non-human agents, such as web crawling software, could navigate, a lot more information could be related much more quickly than now, with the associated epistemological gain. Distributed information systems multiply this potential. If I need to create a digital edition that takes its images from one server and takes its transcriptions from another, and its annotations from yet another, I have to make the interface application on my computer talk to these other computers. And if my computer can, so can other computers.

Thus, we have two epistemological arguments in favour of open distributed digital scholarly editions. The second in particular is an indistinct and opaque artificial intelligence promise at best, colloquially known among developers as the "semantic dream." Even though the dream of a fully digitally networked "apparatus fontium" is beautiful, its pursuit seems not to be a very attractive proposition to textual scholars: most of the essential infrastructure and, veritably, all the appropriately curated content is missing, and creating them requires great effort and technical skills that are not in abundance within the textual scholarship community. The promise of the first argument, that we can further knowledge by leveraging quality of information, is only slightly less opaque, but at least the state of the art in digital infrastructure makes the attainment of this benefit feasible.

Silos and Epistemological Gains

The majority of current digital scholarly editions bring neither of the potential epistemological benefits described above. Most are based on a process of copying, creating or even re-creating all resources in a single digital location (i.e. on one server), forming silos that gather many kinds of different information with different curation and maintenance needs. This situation will not change in light of the promise of networking knowledge via non-human agents that have yet to be designed—the latter of my epistemological arguments above—especially since we have very little inkling of the value of the knowledge that would emerge in this way from distributed information.

The argument about quality of information, and delegation of tasks according to expertise, may actually be convincing for textual scholars. But the incentive for textual scholars to build distributed systems is at best altruistic: creating

webservice-based distributed digital scholarly editions is harder and requires more technical expertise than creating complete and finite websites. As iterated above, the data silo is a cheaper, technically less complicated solution that is less dependent on many external stakeholders, quicker to realize, more adaptable to local needs, and usually has better predictability for deliverables and turnaround.

Does this mean that technically networked information is inevitably, both epistemologically and pragmatically, a non-starter for textual scholarship? That the idea serves no purpose? This remains to be seen. As with so many semantic technologies, the value of distributed information for textual scholarship is more promise than reality, impossible to determine as there are so few real implementations to test-drive. Given the complexities and unclear payoff, it is also not a development that scholars can be expected to lead all by themselves. This is another conundrum: it is on the one hand up to technologists and digital humanists who believe deeply in its promise to demonstrate the value of distributed information resources, but on the other hand the epistemological affordances such networks might create can hardly be left to the technologists to evaluate, as they are usually not textual scholarship experts.

Mirador as an Argument for Distributed Scholarly Resources

So why then would I still maintain, as I said above, that Mirador potentially plays a pivotal role in what may be nothing less than a paradigmatic shift in how we understand, approach and interact with cultural heritage resources? Mirador's strength is in its architectural composition, which a truly lazy reviewer might attack as a mere patchwork of existing code pieces without much added value. But from a networked knowledge perspective, its architecture is precisely the strongest statement it can make, enabling it to be part of a distributed model that would leverage the epistemological benefits of resource quality I argued above. Mirador was built explicitly to do one job and do it very well: viewing digital images. The developers and designers stayed far away from every other temptation. On the functional (or user facing) side of things they provided no image retouching functions, no dedicated transcription possibilities (although the annotation function has been used for simpler transcription tasks), no

metadata editing capabilities, no annotation tools, no print-on-demand-service… no nothing. They just delivered a bare-bones viewing, zooming, panning tool. This is an explicit design choice and by that an explicit assertion on how (scholarly) resources should be networked, namely not by integrating all software and data in one single (server) location, but through lightweight protocols that inform very specialized tools where they can find data and what they should do with it. This strategy allows tools such as Mirador to be completely agnostic as to where some resource is located or how it is produced, served, and maintained, so long as its image data can be requested as input and depicted. In this regard Mirador's architectural makeup can be read as an argument, just as editions (Cerquiglini 1999), interfaces (Andrews and van Zundert 2016), and software code (van Zundert 2016) can be seen as arguments in a wider debate on textual scholarship. Mirador's argument favours distributed digital scholarly resources because it positions itself as a component that fits as a cog in just such a distributed ecosystem. Thus the epistemic argument of Mirador about the digital edition is that a digital edition ought to be a composition of various distributed resources.

Had the developers of Mirador chosen any other strategy, then with every function they added there would have been a tighter integration with other software and stronger demands concerning the form (and possibly location) of data resources— and with every function Mirador would thus have become more an argument in favour of digital scholarly editions as monolithic data silos. In contrast, and quite on purpose, Mirador does not care whether one resource is in Madrid and another is in San Francisco—as the developers themselves explain:

> Users, such as scholars, researchers, students, and the general public, need to compare images hosted in multiple repositories across different institutions. They want a best-in-class experience with deep zoom capabilities, and viewing modalities optimized for single images, books and manuscripts, scrolls, or museum objects. End users want to create and view image annotations, comments, and transcriptions within a single user interface, *regardless of the system in which they were originally created or hosted*. (Sanderson, Snydman, Winget, Albritton, and Cramer 2015; my emphasis)

This makes sense not only from the user's perspective (who does not care where the resource is). It makes sense from a technical point of view too: why duplicate the burden of maintenance and development for all resources? But most saliently: it makes sense from an epistemological point of view: it allows the object of expertise to reside with the expert. It allows the responsibility for the quality of the object to be located in the place best equipped to that end. With a print publication this is harder, as all epistemological objects (transcriptions, structure, contextualization, pictures, index, etc.) are solidified within it. An editor can update the publication, but it takes another expensive print run, and it is unlikely that this will be done in the case of individual changes—the long list of changes (http://vangoghletters.org/vg/updates.html) to the Web-based Van Gogh Letters edition (Jansen, Luijten, and Bakker 2009) testifies to this. In the case of facsimiles, an editor often has to make do with a lower quality photograph of a folio as an illustration (e.g. **Figure 9**). Higher quality can arguably be offered and maintained by an expert in an institution that takes the care
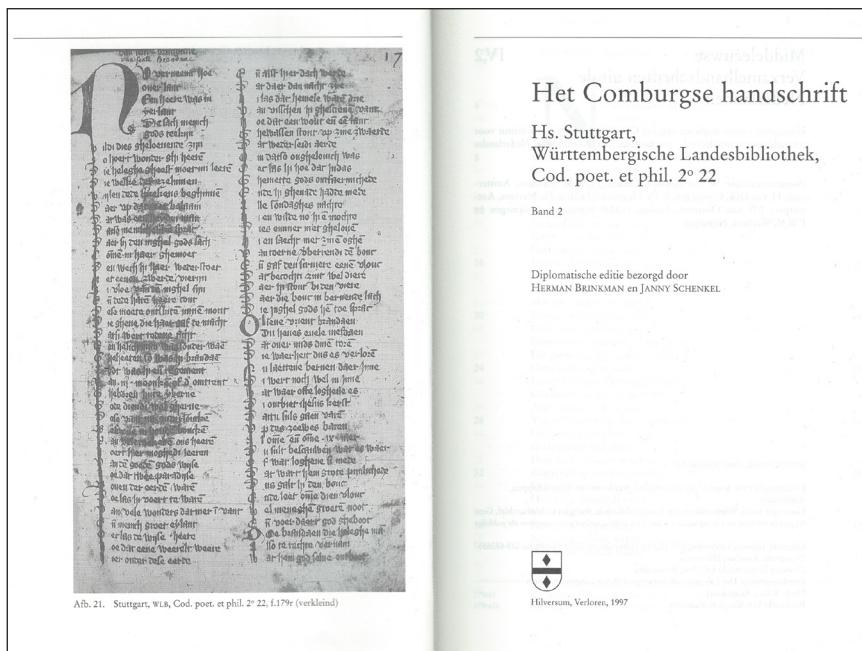
Figure 9: Example of making do with a single photograph. These pages, taken from Brinkman and Schenkel's edition of the Comburg manuscript (Brinkman and Schenkel 1997), show one of the very few reproductions (cut and scaled down) in the print edition. (Image courtesy Verloren Publishers and authors.)

for digital images and its sources to be one of its core tasks. Chances are this will not be the scholar that made the transcriptions and edition. In the case of the edition of the Middle Dutch Comburg manuscript (Brinkman and Schenkel 1997) the repository of the source—the Württembergische Landesbibliothek—did indeed bring high resolution images online, some thirteen years after the print edition was published according to the associated MARC21 information ("Comburger Handschrift – mittelniederländische Sammelhandschrift – Cod.poet.et phil.fol.22" n.d.; "SWB Online-Katalog" n.d.). The facsimiles and the diplomatic transcripts of one of the "flagships" of middle Dutch literature, the Comburg manuscript, lead for the moment a rather unsatisfying—from the scholarly perspective—divorced life. The facsimiles are available on the website of the Würtembergische Landesbibliothek, the full diplomatic transcript only as an offline print edition. Being able to link them up through an architecture as proposed by Mirador would no doubt greatly improve the epistemological value of both.

Mirador as Part of an Ecosystem of Digital Scholarly Resources

Even though it may be reasonable to expect an epistemological benefit from distributed digital scholarly editions, it remains to be seen whether that benefit would actually be realized. The answer is highly dependent on the facility of the technical solution provided. That is: how well and how easily does Mirador let itself be used by scholars and developers alike?

If Mirador is set up in the right way and if the image repository it uses supports the IIIF protocol, then Mirador brings a scholar a long way. IIIF is short for International Image Interoperability Framework (http://iiif.io/). If you want a distributed ecosystem of scholarly resources—that is, the ability to reuse resources no matter where they are actually located—you need some kind of formal language that allows the different services that are resource consumers to know what the resources are, how they are structured, and how they may be used. That sounds high-tech, but in fact the core of it is very social: it amounts to a group of people agreeing on how things will be strictly written, and then adhering to the agreed upon semiotics. If the semiotic signs and rules are rigid, algorithms can process them. The IIIF protocol is one such formal language for the online access of digital images. IIIF was a grassroots development from a community that saw the need for sharing digital image information, and now

seems to be thriving (http://iiif.io/community/#participating-institutions). The various adoptions of the IIIF protocol for specific viewing modalities, such as in the Universal Viewer (http://universalviewer.io) or in Leaflet-IIIF (https://github.com/mejackreed/Leaflet-IIIF) testify to its success. Mirador is one of the kernels in IIIF's developing ecosystem of open source image tools that facilitate the sharing of open access image resources on the web.

For an image repository on the Web to advertise its content according to IIIF rules, it needs to serve a so called manifest file in JSON format (https://www.json.org/) that describes the content and structure of the given resource. The exact semiotics are all documented in detail on the IIIF site (http://iiif.io/api/presentation/2.1/#manifest). The other essential component of an image repository is a server that will stream requested images to the application ("client") that wants to use them. Not any image server will do, however. Again, IIIF compliance is a prerequisite for the server to be part of a distributed network of resources that Mirador asserts. Several of such servers exist "off the shelf" though (cf. http://iiif.io/apps-demos/#image-servers), such as IIPImage (http://iipimage.sourceforge.net/).

Any Mirador viewer can be pointed to a particular manifest by clicking the "Replace Object" option (**Figure 10**). This will let the user choose from a list of pre-selected repositories (**Figure 11**), but a manifest URL can also be keyed in manually
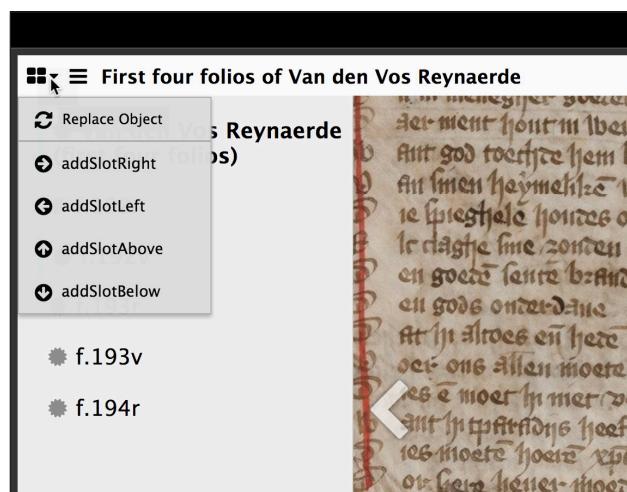
**Figure 10:** Mirador's "Replace Object" function.
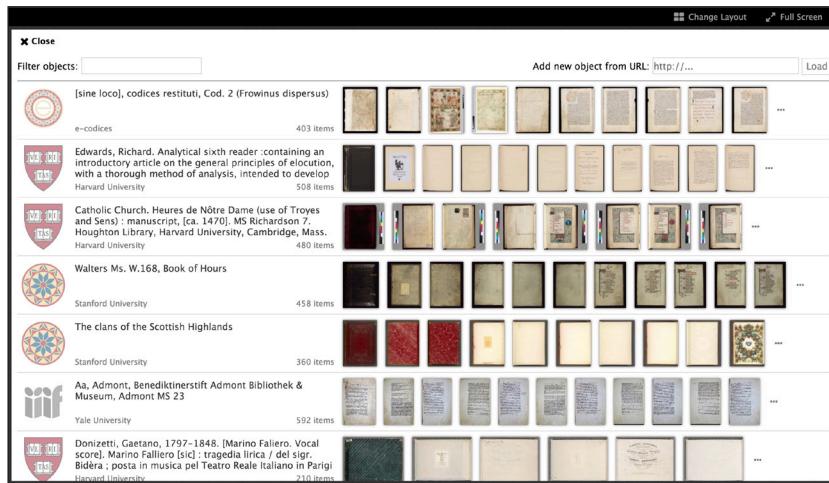
**Figure 11:** An Example of a list with image repositories.
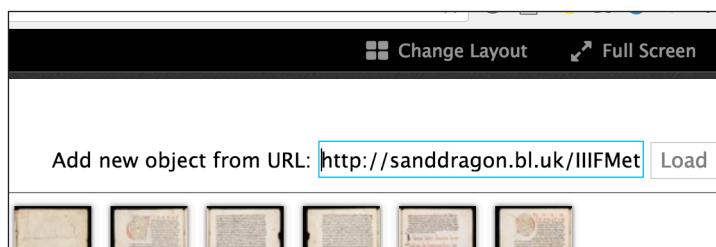
Figure 12: Manually adding an image repository's manifest location.

(**Figure 12**). If Mirador is thus pointed to a location such as http://sanddragon.bl.uk/ IIIFMetadataService/Cotton_MS_Claudius_B_IV.json, it will find all the information it needs to start streaming images to the user—in this case the facsimiles of an incomplete Old English Hexateuch ("Cotton MS Claudius B IV" n.d.). One can gauge from this how distributed Web resources work: all the client (Mirador) really needs to know is the single URL that locates the manifest in the image repository. Because this manifest adheres to the IIIF protocol, both image server and client can operate seamlessly without any geographical or institutional constraints.

Building a Digital Edition with Mirador

Suppose we have a world of distributed scholarly resources: there are repositories of facsimiles, and other repositories serving transcriptions of these facsimile, and yet other repositories may have annotations pertaining to these materials. In a world of

distributed scholarly resources that connect to each other via APIs, speaking certain protocol languages to each other, one should expect facsimiles, transcriptions, and annotations to be different independent resources that are polled and visualized together by a dedicated Web application. Depending on the resources polled, different editions may be created using, to a certain extent, the same resources. This situation is conceptually visualized in **Figure 13** (in comparison to the currently more common monolithic digital scholarly edition).

Suppose a scholar would want to create a digital scholarly edition from such distributed resources. What would it take? To make this more than just a theoretical wish or a thought experiment, I have implemented my own demo of a distributed edition with a tiny sample of images and text (cf. Appendix 1). I mainly wanted to know how hard it would be, because the simpler the work, the likelier that scholars will take to developing digital editions from distributed resources. My experience, however, suggests that one needs to be quite an experienced (web) developer to be able to create the needed web services and the integrating application, i.e. the edition. On a somewhat higher level of overview, what follows describes the implementation of the demo edition.
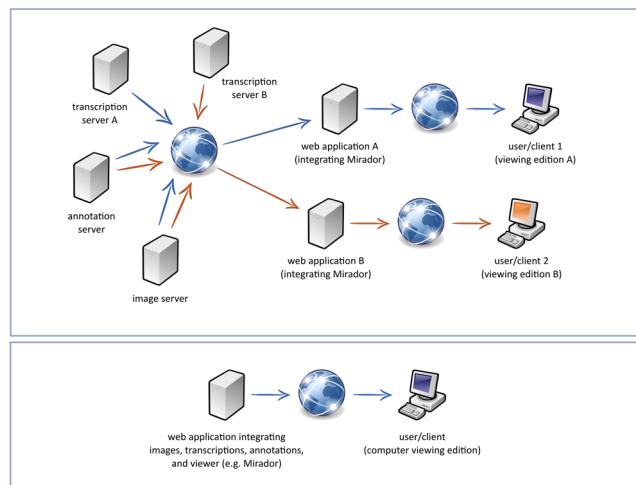
Figure 13: Distributed digital editions varying resources (top) vs. singular integrated monolithic edition (bottom).

Setting up an image server and a Mirador instance is not trivial. One has to know one's way around a Web server such as Apache (https://httpd.apache.org/) and how to run it securely. Installing an image server such as IIPImage (http://iipimage. sourceforge.net/) is relatively straightforward, but having it work properly and securely with Apache involves less than trivial configuration. Apache and IIPImage together form the engine of an image repository. Once both are installed and working together properly, the front page of the image server can be admired (**Figure 14**).

The fuel for the Apache-IIPImage engine is images. These images need to be prepared as so-called pyramidal TIFFs that allow efficient and fast streaming of image information to a client (such as Mirador). Effectively, each image is stored in several sizes in one file to support zooming. Each size is associated with a layer, and each layer is broken up into many small tiles that travel the Internet quickly and easily. To create pyramidal TIFFs, one needs to be comfortable with a tool such as ImageMagick (https://www.imagemagick.org) and commands such as `convert original_0001.tif -define tiff: tile-geometry = 256 × 256 -compress jpeg 'ptif:pyramidal_0001.tif'`. Once the pyramidal images have been created, the image server is able to show us facsimiles (**Figure 15**).

Finally, the manifest file needs to be created. This would usually require a developer to generate it from some database of image information, or by creating a script to query the image files directly (e.g. using exiftool, https://www.sno.phy.queensu.


IIPImage Server

Version 1.0

Project Home Page: http://iipimage.sourceforge.net

**by
Ruven Pillay**
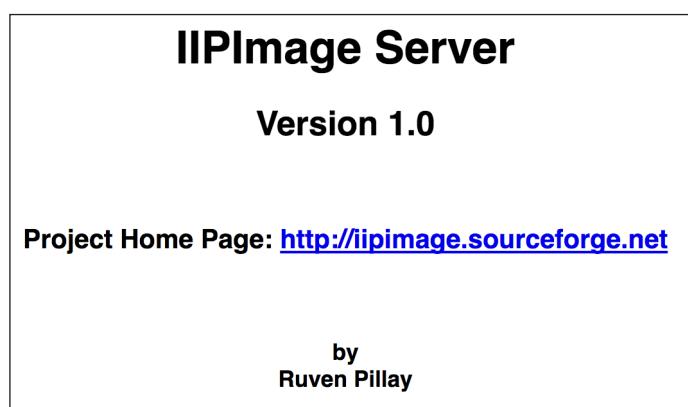
Figure 14: An empty image repository is born.
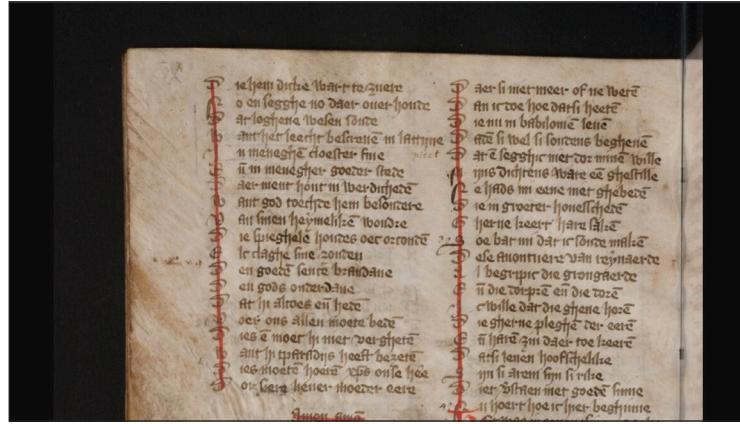
Figure 15: A first facsimile.

Figure 16: Manifest file sample.

ca/~phil/exiftool/). For this demo, given the very few images it would describe, the manifest file was put together by hand. The hardest part of this was understanding the IIIF specification, which is to say, the formal language used to describe the structure and metadata of image collections. The specification is not really complicated, but getting to know it still involves somewhat of a learning curve. There is a tool that is tremendously helpful when building manifests by hand or when trying to familiarize oneself with the IIIF manifest specification, which is the Oxford Manifest Editor (http://iiif.bodleian.ox.ac.uk/manifest-editor). Whether generated or written by hand, the result is a manifest file (**Figure 16**) that can be consulted by any computer running an IIIF client.

At this point there is an image repository that contains, presumably, the facsimiles of the codex that the scholar wants to turn into a digital edition (e.g. something like **Figure 15**). Arguably this part of the work would be delegated to some specialized service or institution. Of course, if no institution already hosts the images one needs, the editor faces the task of convincing some institution to host them, and possibly to fund the related work and maintenance. Alternatively, the editor could create a self-maintained repository as explained here.

Now an actual Web application is needed that uses Mirador to look at the image repository. These days Web application development requires more and more knowledge beyond "plain old HTML and JavaScript." Mirador is no exception, and reusing Mirador's source code while at the same time adhering to more formal and current software development principles requires substantial Web development experience and knowledge. Mirador is developed using the Node.js runtime environment (https://nodejs.org/en/). This means it can actually be run out of the box on Node.js as a server, which is a solution one might opt for. The main reason for choosing Node.js however, seems to be the NPM package manager (https://www.npmjs.com/) that comes along with the environment and which protects developers from the proverbial "version hell"—that is, the conflicts that arise when two components one needs require two different versions of a third component. A 19th century cart wheel will not fit your 21st century Tesla, even though they are just "wheels" and "vehicles". Mirador uses a lot of third-party JavaScript components, and so it needs to carefully check the versions of those components that it combines. NPM is a very convenient way to deal with this problem. However, the consequence is that if you want to use the Mirador source code from its official repository (https://github.com/ProjectMirador/mirador) in the "proper" way, you will first need to download and compile all components and sources that Mirador uses into one mirador.js file using Grunt (https://gruntjs.com/), which is another tool in the Node.js domain. Once this is done, the Mirador demo application provided by the original authors can be deployed by moving Mirador's whole directory into the folder designated by the Apache web server as the source of the files it serves. Of course, there is also a less formal but more convenient and speedier way to start using Mirador. If one does not require substantial changes and adaptations, it is possible to "drop in" a single

pre-compiled mirador.js file, and to refer to this inside a web page's HTML source. Mirador's official Github repository provides pre-compiled versions for this (https://github.com/ProjectMirador/mirador/releases).

At this point we can reach our Mirador instance via any web browser, and we can add the URL of our manifest, which Mirador should use to show us the contents of our image repository (**Figure 17**).

We need a source for our transcriptions too. This involves setting up another server that will provide on request the transcription of a particular page of the codex Mirador is displaying. One could adapt one of the transcription and visualization environments named in the beginning of this article. For my demo I created my own basic transcription server (https://github.com/jorisvanzundert/mirador_review_demo). It uses a Sinatra Web server in the Ruby language (http://sinatrarb.com/; https://www.ruby-lang.org/en/) and serves a TEI-XML file that transcribes a tiny portion of the first facsimile (**Figure 18**; see Appendix 1 for the full source of the textual data). Using the Nokogiri HTML/XML parser in the background (http://www.nokogiri.org/), the same server will use an XSLT stylesheet on request to transform the TEI-XML into HTML, visualizing either a diplomatic (**Figure 19**) or critical version of the transcription.

A server for the images now exists, and we have a server for the transcriptions. But the Mirador client still needs to be told where to find the transcriptions. For the demo I wrote a JavaScript component called "text_viewer" that can request the HTML representation for either the diplomatic transcription, critical transcription,
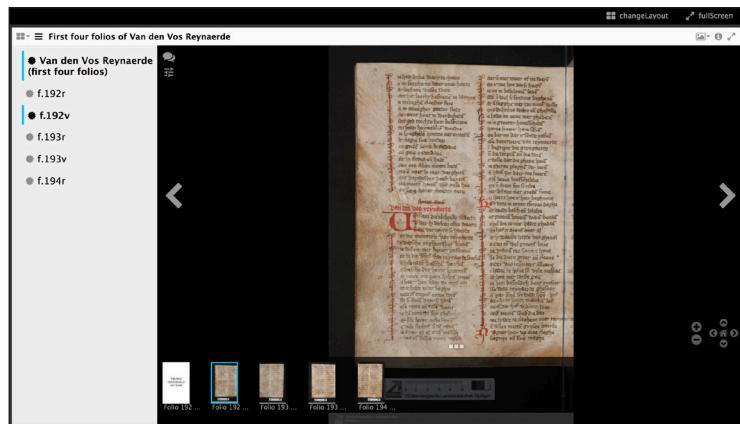
Figure 17: Mirador is up and showing images.

```
    </revisionDesc>
  </teiHeader>
  <text>
  <body>
   <div type="folio" n="192r">
    <div type="part" n="1_0">
     <l>(Transcription of 192r not implemented.)</l>
    </div>
   </div>
   <div type="folio" n="192v">
    <head>
     <title>
      <choice>
       <orig>van den vos reynaerde</orig>
       <reg reason="&rtc;">Van den Vos Reynaerde</reg>
      </choice>
     </title>
    </head>
    <div type="part" n="1_1">
     <l> <c facs="http://localhost:9090/simpleAnnotationStore/annotation/1499201214016">W</c>illem die
      <subst>
       <del>
        <choice>
         <orig>m</orig>
         <reg reason="&rcc;">M</reg>
        </choice>adocke</del>
       <add>vele bouke</add>
      </subst>
     maecte</l>
     <l facs="http://localhost:9090/simpleAnnotationStore/annotation/1499201102798">Daer hi dicken o<choice>
      <abbr>ñ</abbr>
      <expan>mm</expan>
     </choice>e waecte</l>
     <l facs="http://localhost:9090/simpleAnnotationStore/annotation/1499201124745">Hem vernoyde so haerde</l>
     <l facs="http://localhost:9090/simpleAnnotationStore/annotation/1499201200639">Dat die avonture van
      <choice>
       <orig>r</orig>
       <reg reason="&rcc;">R</reg>
      </choice>eynaerde</l>
     <l facs="http://localhost:9090/simpleAnnotationStore/annotation/1499201227082">In
      <choice>
       <orig>d</orig>
       <reg reason="&rcc;">D</reg>
      </choice>ietsche onghemaket blev<choice>
       <abbr>ē</abbr>
       <expan>en</expan>
      </choice>
     </l>
     <l facs="http://localhost:9090/simpleAnnotationStore/annotation/1499201439941">Die
      <app>
       <lem wit="Dycksche">Arnout</lem>
```

Figure 18: The TEI-XML file for the transcription (fragment).

van den vos reynaerde

Willem die vele bouke maecte
Daer hi dicken om̄e waecte
Hem vernoyde so haerde
Dat die avonture van reynaerde
In dietsche onghemaket blevē
Die willem niet hevet vulscrevē

Figure 19: A diplomatic transcription served through the SimpleTranscriptionServer.

or the TEI-XML source from the TranscriptionServer. This component was integrated with the Mirador viewer which results in an application that can show facsimile and transcription together (**Figure 20**; consult Appendix 1 for the source of the text_viewer component).

Along the Seams of Mirador

This is where we start to stumble upon some of the limitations of Mirador. The scholar who studies manuscripts would want a more granular connection between text and image. But as Mirador's developers chose to pursue one task and one task very well, any extras that the scholar might want will have to be added by someone
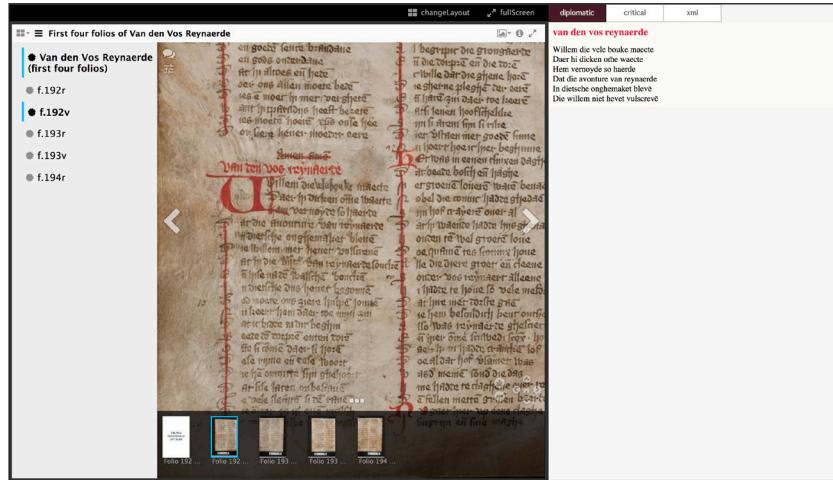
Figure 20: A client combining Mirador viewer for facsimiles and a text viewer for transcription.

with software development capabilities. All this is doable, but it is harder work than what I outlined above, and there I omitted the nitty gritty details of trying a number of unsuccessful solutions that I abandoned in deep heaps of Linux system level error messages that I—being a skilled web application developer but not a very skilled systems admin—could not solve quickly or conveniently.

Up to this point, development consisted of adding whole components together into services that could usefully speak to each other. To realize a more granular linking between text and image, however, we will have to delve into the code of Mirador itself and make some things possible that are not supported by default.

This is, then, the point where we get a feel for the seams of Mirador, for the rough edges of its codework. As much as reusing and wrapping components is a most brilliant strategy to reduce maintenance and reinvention, it has certain disadvantages. Mirador uses code that has been written by many different people, and this shows. Programmers use different styles of coding—and there are many styles (cf. e.g. Croll 2014). It appeared to me that the developers of Mirador are apt JavaScript programmers, which makes for well-written code. Even so, it is not like looking at a Rembrandt or a Van Gogh. It is more like Rembrandt, Van Gogh, and Picasso came together and decided to work concurrently on the same painting. For scholars or programmers wanting to integrate Mirador into their project and

add functionality, this can be a very real difficulty. Another bother, but this may be more of a personal pet peeve, is Mirador's continuous use of object encapsulation through `JQuery.extend()` (https://api.jquery.com/jquery.extend/). This turns Mirador effectively into a God class (https://en.wikipedia.org/wiki/God_object) where everything is connected to everything else, but not everything is necessarily clearly and consistently named. Finding the right hooks and slots to adapt Mirador to your wishes is therefore harder than might have been necessary. This is also not helped by the fact that Mirador's quick-start documentation (http://projectmirador.org/docs/docs/getting-started.html) is very much in a beta phase and that its API documentation is nonexistent (http://projectmirador.org/docs/docs/api-reference.html). Although it is rudimentary, the documentation gives a seasoned web developer just enough hints and insights that she might find her way through. If she wished, this developer could hook into Mirador's code to achieve a more granular linkage between facsimile and transcription. In the demo I wanted to be able to click on any verse in the transcription to cause Mirador to pan and zoom to the corresponding verse on the facsimile. This would seem to me a basic prerequisite of convenience for any digital scholarly edition that ties together medieval text and manuscript facsimile, because either the transcription is used as a reading aid or the facsimile is used to verify the correctness of the transcription. Such linking requires some way of relating a particular TEI l-element (i.e. line element, http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-l.html) to a particular area of the facsimile. Here we come upon a rough seam in the IIIF protocol that Mirador is using.

According to the IIIF specification, there are actually several ways to achieve a more granular linking between text and image. One is the ability to define ranges—these may indicate a range of pages, for instance, or a particular area on a page (http://iiif.io/api/presentation/2.1/#range). Another way is to use a segment selector in a URI (http://iiif.io/api/presentation/2.1/#segments). Neither solution is very satisfying, however. For one thing, the IIIF specification is still an evolving specification, and the ranges model is a good example of its current volatility. IIIF community discussion around ranges led to the deprecation of the current specification for ranges, which is to be fully replaced in IIIF version 3.0 (https://github.com/IIIF/api/issues/1070).

More importantly, both solutions assume that knowledge about the transcription is part of the description of the facsimile. From the point of view of decoupled distributed scholarly resources maintained in a context of specific expertise, this is unsatisfying because it establishes a strong coupling between the facsimile resource and text resource. In keeping with the idea that services should be as content-agnostic as possible, it would be preferable that any Mirador application, similar to how it reaches out to an image repository, would reach out to an external service that on request would provide just enough information for it to link specific parts of the transcription with specific parts of the facsimile.

Fortunately, the authors of the IIIF specification decided to follow the web annotation recommendations by the W3C that provide exactly this type of model (Cole 2017). The web annotation model became an official W3C recommendation on 23 February 2017 after years of preparation by the W3C Web Annotation Working Group (https://www.w3.org/annotation/), which derived from another grassroots initiative called the Open Annotation Collaboration (http://www.openannotation.org/). It is this Web Annotation Model that allows us to call another independent service into action that exposes annotations to the Web that are independent of the description of that facsimile by the image resource. Mirador can then retrieve annotations on a facsimile from this service. Conveniently—for the demo in any case—there exists a ready-made SimpleAnnotationServer that will act as just such an independent resource for annotations (https://github.com/glenrobson/SimpleAnnotationServer). This service can provide information about annotations available for the facsimile to both my self-rolled SimpleTranscriptionServer and to Mirador. Mirador is able to depict these out of the box (see **Figure 21**). However, to enable the user to click on a particular part of the transcription and to have Mirador then pan and zoom to the corresponding part of the facsimile, we need to hack a few lines of code deep within Mirador's innards. Mirador encapsulates its own event handler mechanism that enables it to publish events and to subscribe to events of other components. For example, if a user clicks on a verse, the custom-made text_viewer component publishes an event called `request_fit_bounds`, which includes information about the location of the image that was clicked.
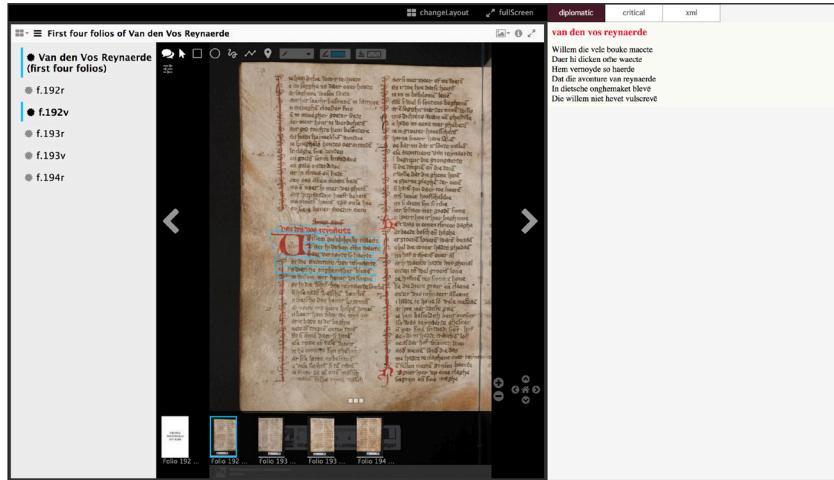
Figure 21: Mirador showing annotations, annotated areas with blue borders.

After we have modified its code a bit, Mirador can listen for these events and, when one occurs, pan and zoom to the verses that correspond to the location given in `request_fit_bounds`. Clicking on the first character of the transcription, for instance, zooms to the enlarged initial just to make the case (**Figure 22**).

The fact that IIIF relies on the Web Annotation model of the W3C is fortunate in more than one sense. It ties into my contention that knowledge quality is best served when it resides with exactly the expertise it needs. That in turn serves to keep knowledge within repositories to the bare minimum needed to serve a very specific purpose, and with this comes efficient maintenance and other practical benefits. But it is also fortunate in the sense that, this way, the IIIF specification runs less of a risk of bloating. The same phenomenon that is a risk for integrated infrastructures (i.e. that they topple under the maintenance of ever more tools and data being integrated) is a risk for protocols and standards too: their authors may be tempted to expand their coverage and expressiveness forevermore. Signs of this kind of bloating may be found in the resource structure specification of IIIF, that deals with what is actually on the images, how they form a collection, etc. (http://iiif.io/api/presentation/2.1/#resource-structure). The lure to over-specify is a real pitfall, but judging from the technical community discussions (https://github.com/IIIF/api/issues/1070) and deprecation warnings (http://iiif.io/api/presentation/2.1/#collection), it looks like the community
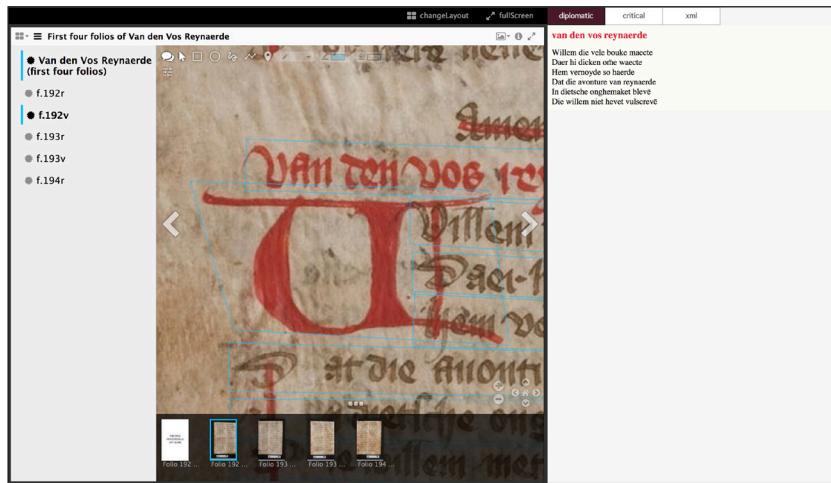
Figure 22: Mirador zoomed in on the enlarged initial of the manuscript.

is veering towards a conservative policy in adding specifications, which would be a tremendously good thing. Literally anything can be depicted on an image, so it is essential to carefully avoid describing picture content in order to keep a lean and effective interoperability protocol, which is what IIIF wants to be. Defining what is on the image is better left to more specific community standards and protocols. In the case of manuscripts and codices, one can indeed imagine a very productive "handoff" of content description according to the TEI model. At the same time that this makes sense, it is also hard. Textual scholars, and especially the digital adepts among them, have upheld a naive notion of unproblematic separation of materiality and textual content for a long time. DeRose's publication on the OHCO (Ordered Hierarchy of Content Objects) can be taken as a convenient temporal marker for the emergence of this attitude (DeRose, Durand, Mylonas, and Renear 1990). In reality, text and materiality are deeply intertwined and their separation is not unproblematic at all (Galey 2010, 110–4). The trickeries of such an illusory separation almost immediately surface when one starts working with Mirador and codices. Mirador sports a "bookView" function (see **Figure 23**), which presents every two images as pages that face each other in a book. The "bookView" function however, assumes that a series of images is 1-to-1 related with a series of images of consecutive individual page sides, and that the first in a series of images is a depiction of a right-hand side page face. Neither needs to be true, even if these assumptions are in line with what is more or less a general convention in
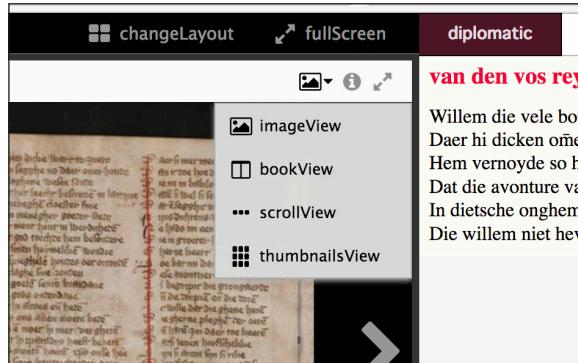
Figure 23: Mirador's "bookView" function.

the modern West. In the case of the demo presented, the chosen example text starts midway left column on a verso of a folio. By default, the bookView function would depict it as a right-hand side page. The only way around this at the moment is to insert a page intentionally left blank (cf. **Figures 24** and **25**). The IIIF specification that Mirador adheres to has little way of expressing constructs that intersect material and textual dimensions. This is a case where Mirador's design makes implicit assumptions that are not backed by its IIIF model. Keeping with the technically sensible idea of separation of concerns, a book view function ought only to be implemented if there is actually a model (e.g. TEI) that informs it about the relation between text and pages, and between pages and images. It is not for me to claim that Mirador's developers were unaware of this—I simply do not know. The problem may have been that the specification is still volatile exactly on this issue. But this does show how easy it is to misconstrue the reach of a specification—with, I think, harmful epistemological consequences. The reader may point out that on the level of world history, these consequences are not actually all that harmful. Fair enough. But taking digital textual scholarship seriously can hardly be reconciled with leaving potential confusion over basic documentary information such as which pages were right and left.

In the case of manuscript images and scholarly transcriptions, IIIF and TEI have much to gain from each other. But it remains to be seen how they should be aligned or connected. This is an issue that has already seen some recent discussion on the TEI mailing list (cf. Stutzmann 2017). In the demo, I challenged myself to make Mirador pan and zoom towards a particular area that is annotated with a particular part of
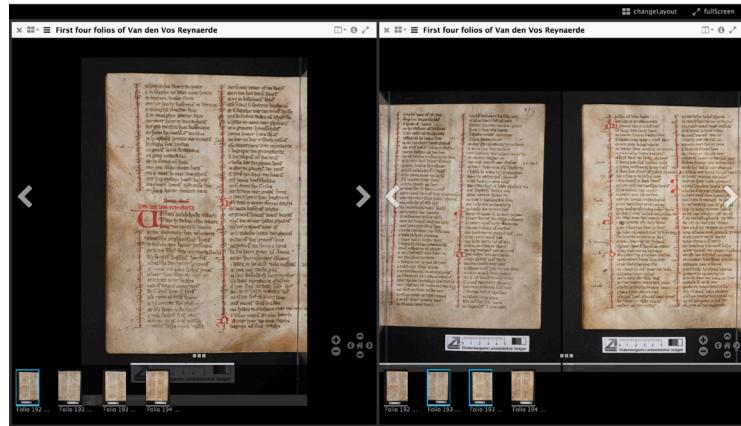
Figure 24: Mirador's unmitigated book view, suggesting that this text starts on a recto.
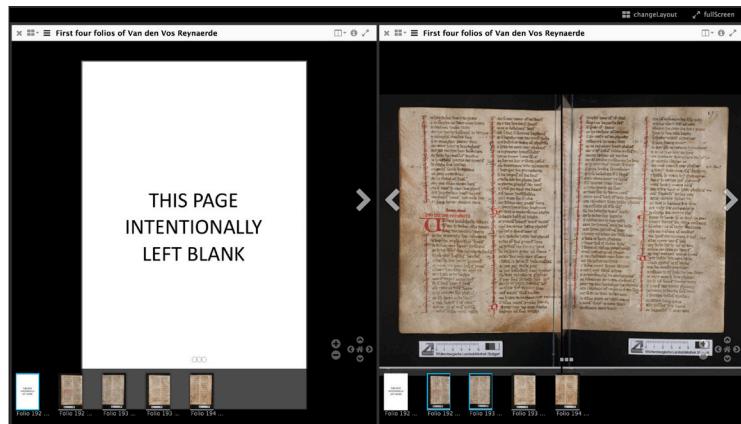
Figure 25: Mitigating Mirador's book view by inserting a blank page.

the transcription. This was also the use case that brought the TEI-L discussion to life: how to implement sub-page granular referencing between text (transcription) and image (page). From the point of view of distributed and decoupled scholarly resources, one does not want specific knowledge of the text description integrated within the image (IIIF) description, nor, vice versa, does one want knowledge about the image description tightly integrated into the text (TEI) description. Thus, neither IIIF's proposal to link directly into a TEI file by using XPointer and XPath (http://iiif. io/api/presentation/2.1/#segments) queries:

```json
{
 "@context": "http://iiif.io/api/presentation/2/context.json",
 "@id": "http://localhost:9999/annotations/annotation/anno1",
 "@type": "oa:Annotation",
 "motivation": "sc:painting",
 "resource":{
   "@id":
"http://localhost:9099/reynaert/diplomatic/folio/192v#xpointerfolio/19
2v#xpointer(tei:text/tei:body/tei:div[@type='folio']/tei:div[@
type='part' and @n='1_1']/tei:l/tei:c)",
  "@type": "dctypes:Text",
  "format": "application/tei+xml"
 },
 "on":
 "http://localhost:9999/reynaert_fragment/folios/folio_192v#x
 ywh=100,100,500,300"
 }
```

nor the solution proposed on the TEI list (cf. Holmes 2017), to strongly couple parts of transcriptions to segments (areas) of images using, for instance, a `facs` attribute:

```xml
<div type="part" n="1_1">
    <l><c
facs="http://localhost:9999/reynaert_fragment/folios/folio_192v#xywh=1
00,100,500,300">W</c>illem die
      <subst>
        <del>
          <choice>
            <orig>m</orig>
            <reg reason="&rcc;">M</reg>
          </choice>adocke</del>
        <add>vele bouke</add>
```

is really satisfying. Again, the reason is that editors of scholarly texts are best not bothered with specific image-related description, and a protocol or standard ought not to push this highly specialized knowledge on them. This is of course not to say that it

is forbidden ground for the textual scholar, but if she wants that knowledge she ought to find it in the designated place, and should not be bothered by it out of necessity.

The schemes above, moreover, effectively preclude competing transcriptions. Suppose you have two competing transcriptions for the same facsimile. With the strong coupling of the transcription fragment inside the image segment (`book1/canvas/p1#xywh=0,0,600,900`) in the IIIF scheme, a viewing client like Mirador has no choice: the image description dictates that the client should go look for one specific transcription (that of the very specific XPointer denoted in the segment definition). This type of strong integration goes exactly against the nature of distributed resources and nullifies the ability to discover distributed knowledge. If there are multiple competing transcriptions for one particular facsimile, then a viewer for that facsimile should be able to discover any or all of the transcriptions. The strong coupling above forces this work of discovery onto the creator and/or maintainer of the facsimile image, which is to say, a person whose immediate interest and expertise is probably not geared to that task. Instead, and in the interest of epistemological gain, we ought to register the transcription with an additional independent service. A client like Mirador can in that case just ask from such a service: "Is there any transcription for this particular facsimile?" The service will then answer with the appropriate resources, and if there are competing transcriptions, the viewer can choose one or present them as alternatives. Introducing such an intermediate service is called "adding an indirection," or "making resources dereferenceable," to put it in stark information technology terminology.

To my knowledge there is no community consensus about a formal protocol to support this type of service. It could be tremendously productive if both the IIIF and TEI communities were to enter in a dialectic on that topic. For the moment this type of behavior can be mimicked at best by utilizing the Open Annotation schemes that IIIF adheres to, as demonstrated by the Mirador-based application presented.

Conclusion: The Risks to Mirador's Distributed Worldview

Where do we stand after a long journey from Mirador along IT architecture, monoliths and epistemology, to distributed knowledge and the construction of a scholarly edition demonstrator from distributed resources? My conclusion is

that things look pretty bleak with regard to the potential success of distributed resources in scholarship. Not, of course, because distributed resources are a bad idea. In fact, I would contend that from a scholarly point of view a distributed architecture is the only IT information and knowledge architecture that makes sense. It fits better with the tenets of scholarship, which values multiple perspectives and intersubjective interpretation. But IT infrastructure is often overlooked as both a metaphor for and a formative agent of epistemological construction. That it can be a normative epistemological means is—I would argue—hardly recognized in scholarship. Since it is perceived as pretty much unrelated to the core of their epistemology, scholars are thus unlikely to meddle much with the architecture of their IT infrastructure. As argued, there should be a modest epistemological gain in distributed resources. However, as this is currently a technological promise at best, it is only a weak argument in favour of distributed digital scholarly resources.

The deep involvement of scholars with the requirements specification for Mirador and IIIF have no doubt contributed to their respective success stories so far. My worry, however, is with the integration of these technologies in existing institutional contexts. Mirador makes a very strong statement for distributed architecture and the connection of distributed knowledge and expertise. But this statement is in code, in technical architecture, and in the technical specification of a protocol. As such, it currently can only be really understood at all levels by software developers, or scholars who are skilled developers too. As my own demo shows, a high level of IT expertise is needed to create scholarly tools using Mirador as a component. Of course, this is not to suggest that developers do not listen carefully to scholars formulating requirements. Mirador is obviously an excellent case in this regard, for the developers understood very well the need for scholars to compare images from different remote sources. But this does not mean the distributed architectural vision carries over automatically to local contexts. As I have argued, distributed knowledge is not part of a general scholarly worldview, rather, distributed architecture only emerges as relevant to scholars in the very specific case where they cannot easily obtain a very specific resource. The generalized case of having all digital objects

available as distributed resources remains an abstract idea for them, and moreover an idea whose realization would require tremendous effort while having little concrete epistemological appeal. It is therefore questionable whether the majority of requirements put to developers by scholars when integrating technologies like Mirador and IIIF in a local institutional context will actually push them towards a distributed architecture. I am pessimistic in this respect, and I think it is far more likely that both developer and scholar convenience will favour local repositories and local integrated tools connecting to those local repositories: linking cross-institutional distributed resources requires a lot of overhead in meetings, discussions and collaboration.

In the case of IIIF, as in many other cases, there may also be a discrepancy between who decides on the architecture of the technology and who is assumed to reap the epistemological benefits of it. The visitors and speakers lists of the 2017 IIIF conference at the Vatican reveals a large majority of technologists, and a small minority of scholars (https://2017iiifconferencethevatican.sched.com/directory/speakers). Some well-known names in digital scholarship (Jeffrey Witt, Peter Robinson, John Bryant, Ben Brumfield, Frederik Kaplan) are represented, which is good, but that group should broaden and diversify if it is to avert the next futile technology push. Mirador and IIIF may turn out to be typical examples of technologies that came too early. The scholarship community may not be well-versed enough, and is certainly not yet fluent enough, in computer and IT architecture languages to fully appreciate what distributed resources have to offer and how they constitute a different worldview from monolithic software solutions.

While all those risks of development, mutual understanding, and adoption can be mitigated, there still remain some caveats purely on the technological plane. Authentication, for one, can be a nightmare and the sudden death of any well-argued infrastructure. Barring that, there is still the CAP theorem to contend with (see https://en.wikipedia.org/wiki/CAP_theorem) if distributed scholarly resources were indeed to take off. That said, technical issues are usually the easiest to solve in the case of sociotechnical systems.

Somewhere in between the social and the technical is the question of the likelihood of Mirador's, or particularly IIIF's, adoption by existing repositories. Notwithstanding IIIF's thriving community, it remains to be seen whether

repositories that invested heavily in other technologies such as the DFG Viewer (http://dfg-viewer.de/) adopted by *inter alia* the Württembergische Landesbibliothek (http://www.wlb-stuttgart.de/) and the Universitäts- und Landesbibliothek Münster (https://www.ulb.uni-muenster.de/), will be inclined to support yet another protocol in addition. As so often, the technology is not a showstopper in this case, but institutional politics, development capacity, funding, and policies may very well be. Still, in all, I prefer to think that Mirador got most things exactly right. The choice to limit the viewer to what is the bare minimum of functional essentials, built from reused components and software, is certainly wise. Most of all, the developers successfully avoided bloating the tool under the pressure of feature requests. Hopefully the same will be true of IIIF. Less is more; the leaner the specification, the easier the adoption.

Additional File

The Additional File for this article can be found as follows:

· **Appendix 1.** Note on Running the Mirador Demo. DOI: https://doi.org/10.16995/dm.78.s1

Competing Interests

The author has no competing interests to declare.

References

Andrews, Tara L., and **Joris J. van Zundert.** 2016. "What Are You Trying to Say? The Interface as an Integral Element of Argument." In *Digital Scholarly Editions as Interfaces: Abstracts and Programme*, 29–30. Graz: Centre for Information Modelling – Graz University. Accessed July 5, 2018. https://static.uni-graz.at/fileadmin/gewi-zentren/Informationsmodellierung/PDF/dse-interfaces_BoA21092016.pdf.

Berners-Lee, Tim. 1995. "Hypertext and Our Collective Destiny." Transcript of a talk. Berners-Lee: Talk at Bush Symposium, Cambridge, MA, October 12. Accessed July 5, 2018. http://www.w3.org/Talks/9510_Bush/Talk.html.

Boot, Peter, and **Joris van Zundert.** 2011. "The Digital Edition 2.0 and The Digital Library: Services, Not Resources." *Bibliothek und Wissenschaft* 44: 141–52.

Borgman, Christine L. 2015. *Big Data, Little Data, No Data: Scholarship in the Networked World.* Cambridge: MIT Press.

Bouwman, André, and **Bart Besamusca.** 2009. *Of Reynaert the Fox: Text and Facing Translation of the Middle Dutch Beast Epic Van Den Vos Reynaerde.* Translated by Thea Summerfield. Amsterdam: Amsterdam University Press. Accessed July 5, 2018. http://www.oapen.org/search?identifier=340003. DOI: https://doi.org/10.5117/9789089640246

Brinkman, Herman, and **Janny Schenkel.** (eds.) 1997. *Het Comburgse handschrift: Hs. Stuttgart, Württembergische Landesbibliothek, Cod. poet. et phil. 2° 22* 2. Middeleeuwse Verzamelhandschriften uit de Nederlanden 4. Hilversum: Verloren.

Casey, John. 2012. "Iron Manning, Again." *The Non Sequitor* (blog), September 4, 2012. Accessed July 5, 2018. http://thenonsequitur.com/?p=3527.

Cerquiglini, Bernard. 1999. *In Praise of the Variant: A Critical History of Philology.* Baltimore: The Johns Hopkins University Press.

Cole, Timothy. 2017. "Making It Easier to Share Annotations on the Web." *W3C Blog* (blog), February 23, 2017. Accessed July 5, 2018. https://www.w3.org/blog/2017/02/making-it-easier-to-share-annotations-on-the-web/.

"Comburger Handschrift – mittelniederländische Sammelhandschrift – Cod.poet.et phil.fol.22." n.d. *Württembergische Landesbibliothek Stuttgart.* Accessed February 6, 2018. http://digital.wlb-stuttgart.de/purl/bsz323970265.

"Cotton MS., Claudius B IV." n.d. *British Library: Digitised Manuscripts.* Accessed February 6, 2018. http://www.bl.uk/manuscripts/FullDisplay.aspx?ref=Cotton_MS_Claudius_B_IV.

Croll, Angus. 2014. *If Hemmingway Wrote JavaScript.* San Francisco: No Starch Press.

DeRose, Steven J., David G. Durand, Elli Mylonas, and **Allen H. Renear.** 1990. "What Is Text, Really?" *Journal of Computing in Higher Education* 1(2): 3–26. DOI: https://doi.org/10.1007/BF02941632

Dombrowski, Quinn. 2014. "What Ever Happened to Project Bamboo?" *Literary and Linguistic Computing* 29(3): 326–39. https://doi.org/10.1093/llc/fqu026. DOI: https://doi.org/10.1093/llc/fqu026

Fagin, Ronald, Joseph Y. Halpern, Yoram Moses, and **Moshe Y. Vardi.** 1995. *Reasoning About Knowledge.* Cambridge: MIT Press.

Fecher, Benedikt, Sascha Friesike, and **Marcel Hebing.** 2015. "What Drives Academic Data Sharing?" *PLoS ONE* 10(2): e0118053. DOI: https://doi.org/10.1371/journal.pone.0118053

Fischer, Franz. 2017. "Digital Corpora and Scholarly Editions of Latin Texts: Features and Requirements of Textual Criticism." *Speculum* 92(S1): S265–S287. DOI: https://doi.org/10.1086/693823

Franzini, Greta, Peter Andorfer, and **Ksenia Zaytseva.** 2016. "Catalogue of Digital Editions." Accessed July 9, 2018. https://dig-ed-cat.acdh.oeaw.ac.at/.

Galey, Alan. 2010. "The Human Presence in Digital Artifacts." In *Text and Genre in Reconstruction: Effects of Digitalization on Ideas, Behaviors, Products and Institutions*, Willard McCarty (ed.), 93–118. Cambridge (UK): Open Book Publishers. Accessed July 5, 2018. http://individual.utoronto.ca/alangaley/files/publications/Galey_Human.pdf. DOI: https://doi.org/10.11647/OBP.0008.04

Gitelman, Lisa. (ed.) 2013. *"Raw Data" Is an Oxymoron.* Cambridge: MIT Press.

Goldman, Alvin, and **Thomas Blanchard.** 2016. "Social Epistemology." In *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.). Accessed July 5, 2018. https://plato.stanford.edu/archives/win2016/entries/epistemology-social/.

Holmes, Martin. 2017. "Re: IIIF and Facs." *TEI-L Archives.* June 18. Accessed July 5, 2018. https://listserv.brown.edu/archives/cgi-bin/wa?A2=TEI-L;d0eacf69.1706.

Jansen, Leo, Hans Luijten, and **Nienke Bakker.** (eds.) 2009. *Vincent van Gogh: The Letters.* Amsterdam: Amsterdam University Press. Accessed July 5, 2018. http://www.vangoghletters.org/.

Kiernan, Kevin, Alex Dekhtyar, Jerzy W. Jaromzcyk, and **Dorothy Carr Porter.** 2004. "Edition Production Technology (EPT) and the ARCHway Project." *Digidult. Info.* August.

McGann, Jerome. 2015. "Truth and Method: Humanities Scholarship as a Science of Exceptions." *Interdisciplinary Science Review* 40(2): 204–18. DOI: https://doi.org/10.1179/0308018815Z.000000000113

Nyhan, Julianne, and **Andrew Flinn.** 2016. *Computation and the Humanities: Towards an Oral History of Digital Humanities*. Springer Series on Cultural Computing. Cham: Springer Open. http://www.springer.com/gp/book/9783319201696.

Porter, Dot. 2013. "Medievalists and the Scholarly Digital Edition." *Scholarly Editing: The Annual of the Association for Documentary Editing* 34. Accessed July 5, 2018. http://www.scholarlyediting.org/2013/essays/essay.porter.html.

Prescott, Andrew. 2016. "Beyond the Digital Humanities Center: The Admistrative Landscapes of the Digital Humanities." In *A New Companion to Digital Humanities*, Susan Scheibman, Ray Siemens, and John Unsworth (eds.), 461–75. Malden, Oxford: John Wiley & Sons, Ltd. Accessed July 5, 2018. http://onlinelibrary.wiley.com/doi/10.1002/9781118680605.ch32/summary.

Robinson, Peter. 2004. "Where We Are with Electronic Scholarly Editions, and Where We Want to Be." March 24. Accessed July 5, 2018. http://computerphilologie.uni-muenchen.de/jg03/robinson.html.

Robinson, Peter. 2013. "Why Digital Humanists Should Get out of Textual Scholarship. And If They Don't, Why We Textual Scholars Should Throw Them Out." *Scholarly Digital Editions* (blog), July 29. Accessed July 5, 2018. http://scholarlydigitaleditions.blogspot.nl/2013/07/why-digital-humanists-should-get-out-of.html.

Robinson, Peter. 2016. "The Digital Revolution in Scholarly Editing." In *Ars Edendi Lecture Series*, Barbara Crostini, Gunilla Iversen, and Brian Jensen (eds.) IV: 181–207. Stockholm: Stockholm University Press. DOI: https://doi.org/10.16993/baj

Sanderson, Rob, Stuart Snydman, Drew Winget, Ben Albritton, and **Tom Cramer.** 2015. "Mirador: A Cross-Repository Image Comparison and Annotation Platform." *Paper presented at OR2015| 10th International Conference on Open Repositories*. Indianapolis, IN, June 9. Accessed July 9, 2018. http://program.or2015.net/sanderson-mirador-226.pdf.

Schreibman, Susan, Laura Mandell, and **Stephen Olsen.** 2011. "Introduction". *Profession*, 123–201. DOI: https://doi.org/10.1632/prof.2011.2011.1.123

Siemens, Ray, Constance Crompton, Daniel Powell, and **Alyssa Arbuckle.** 2016. "Building A Social Edition of the Devonshire Manuscript." In *Digital Scholarly*

Editing: Theories and Practices, Elena Pierazzo, and Matthew James Driscoll (eds.), 37–60. Cambridge: Open Book Publishers. Accessed July 5, 2018. http://www.openbookpublishers.com/reader/483.   DOI: https://doi.org/10.11647/OBP.0095.08

Sondheim, Daniel, Geoffrey Rockwell, Stan Ruecker, Mihaela Ilovan, Luciano Frizzera, and **Jennifer Windsor.** 2016. "Scholarly Editions in Print and on the Screen: A Theoretical Comparison." *Digital Studies/Le Champ Numérique* 6. DOI: https://doi.org/10.16995/dscn.14

Stutzmann, Dominique. 2017. "Re: IIIF and Facs (and TEI)." *TEI-L Archives.* June 27. Accessed July 5, 2018. https://listserv.brown.edu/archives/cgi-bin/wa?A2=TEI-L;205cb507.1706.

"SWB Online-Katalog." n.d. *SWB-Online Katalog.* Accessed February 6, 2018. http://swb.bsz-bw.de/DB=2.1/PPNSET?PPN=323970265&INDEXSET=1/PRS=MARC21.

Thiruvathukal, George, Steven Jones, and **Peter Shillingsburg.** 2010. "E-Carrel: An Environment for Collaborative Textual Scholarship." *Journal of the Chicago Colloquium on Digital Humanities and Computer Science* 1(2): 1–12. https://ecommons.luc.edu/cs_facpubs/1/.

Timpanaro, Sebastiano. 2005. *The Genesis of Lachmann's Method.* Translated by Glenn W. Most. Chicago, London: University of Chicago Press. Accessed July 5, 2018. http://www.loc.gov/catdir/toc/ecip0513/2005015897.html.

Van Zundert, Joris J. 2016. "Author, Editor, Engineer — Code & the Rewriting of Authorship in Scholarly Editing." *Interdisciplinary Science Reviews* 40(4): 349–75. DOI: https://doi.org/10.1080/03080188.2016.1165453