



Building Platforms For Historians: Making Data Findable

A Critical Reflection On Some German Initiatives

Lilian Landes

Bavarian State Library | Centre for Electronic Publishing (ZEP)
Kaulbachstrasse 19 | 80539 Munich | Germany
landes@bsb-muenchen.de

Abstract: Each discipline and subject traditionally has specific needs regarding digital work and research environments. Still, essential thematic complexes which are especially important in terms of the conception and establishment of scholarly platforms providing content in Open Access, can be named: Based on specific examples from the field of history in Germany, the article considers four categories: content selection policy, digitalisation, content genres and contextualisation. With the development of user requirements over the last 10 years, the challenges involved in the selection of provided content, the awareness of copyright issues or the linking and spreading of provided content are also changing.

Keywords: Humanities, Open Access, History, scholarly communication

Reading habits, work tools and communicative mechanisms are changing amongst scholars in the age of increasing digitisation. The natural sciences were pioneers in developing adequate instruments, but several initiatives have since also rooted in the humanities (especially in the historical sciences) and have long since passed the stage of testing. As more come into existence they show a significant transformation such as the move away from result-oriented towards a process-oriented supply of scholarly material. Many permanently successful projects are based on the commitment of individual persons or small networks. In Germany, a kind of 'e-project family' has emerged during the last 10 years around Gudrun Gersmann, for instance, the director of the [German Historical Institute](#) in Paris, and professor for Early Modern History at the University of Cologne. Some of the most important online tools (e-journals, platforms) for German historians, such as [recensio.net](#), [perspectivia.net](#), [sehpunkte](#) and [historicum.net](#) are based (partially also) on her initiative.¹ All of them are committed to the basic idea of Open Access.

¹ A current overview of "digital history" in Germany is provided by Thomas Meyer: 'Virtuelle Forschungsumgebungen in der Geschichtswissenschaft – Lösungsansätze und Perspektiven', *LIBREAS. Library Ideas*, 7, 1 (18), 2011, http://www.ib.hu-berlin.de/~libreas/libreas_neu/ausgabe18/texte/05meyer.htm [June 15, 2011]. Cf. also: Winfried Enderle: 'Geschichtswissenschaft, Fachinformation und das Internet', *eForum*, 3/4, 2001, p. 1-9, http://www.eforum-zeitgeschichte.at/3_01a7.html [June 21, 2011].

The fundamental significance of digital source collections is no longer widely questioned. The advantages of e-collections and e-journals (e.g. faster distribution, worldwide availability, targeted contextualisation of content) are obvious. The related challenges (long-term preservation, international copyright issues, especially with regard to Open Access etc.) are apparent in all of the projects mentioned.

In this article I will look at certain aspects of different e-concepts, and integrate them into four more general contexts and principle questions: first, the issue of *content selection policy*: How does recensio.net deal with this question during its initial phase? Secondly *digitisation*: What general difficulties are there? Thirdly: How 'attractive' are different content genres for an Open Access publication? And finally, *contextualisation*: Which specific measures can be helpful in terms of making scholarly material easier to find for the target audience?

1 Content Selection Policy

In Mykonos in June 2010, a fairly clear content selection policy for EUscreen was agreed upon with 14 historical topics and comparative virtual exhibitions constituting the thematic outline. A precisely outlined content selection policy conveying a clear message to both users and future content providers of a platform (meaning cooperating partners) is vital for any future success.

1. Profile

One of the most destructive factors during the start-off-phase of a platform for historians is any kind of arbitrariness regarding the content selection. Even worse is the absence of a distinct profile on what to integrate, which often results from a lack of ideas about the target group and their interests.

2. 13 Years Ago

Traditional formats born at the end of the 1990s often started with concepts that had a broad thematical focus. historicum.net for example was launched 13 years ago as 'Server Frühe Neuzeit' (Server Early Modern History). It was one of the first web-platforms for German historians, with a broad selection of provided text genres, covering different periods of time and offering various types of source and secondary material. It rapidly became one of the most popular web-meeting points for young history students and lecturers. At that time, a historian was thankful for almost any kind of information which was easily available online.

3. New Web

Today these platforms often suffer from constantly growing competitors, meaning more and more competing services, which are becoming increasingly more specific and cross-linked. Primarily, this is the case with successful Web 2.0 formats like Wikipedia, which grow much faster and can slowly eliminate the scholar's prejudice about the quality of their content.² Another important aspect is the fact that scholars contributing articles to a platform like historicum.net have less and less time, whereas more and more people are joining Web 2.0 concepts like Wikipedia, thus enhancing their quality by increasing the rate of competition.

Today, the concept of any new platform should be clear right from the start, past experiences have highlighted the advantages of setting tight limits concerning the character of the content, and – of course – the advantage of being very selective concerning its quality, especially in the beginning. At the same time, there has to be a critical amount of content when launching a platform. Living up to both of these requirements can be a balancing act.

4. Recensio.net

An example of the process of creating and launching a new platform nowadays is recensio.net – review platform for European History, which is based at the Bavarian State Library and funded by the German Research Foundation (DFG). In January 2011, recensio.net went online. It is a Europe-wide Open Access platform for reviews of historical literature: On the one hand, it brings together 'classical' reviews, meaning those published in historical journals and on the other, it establishes a Web 2.0 based form of scholarly reviewing. That means that

² Cf. Peter Haber, Jan Hodel, J., 'Geschichtswissenschaft und Web 2.0. Eine Dokumentation'. The hist.net Working Paper Series, 2, p. 22-23; Daniela Pscheida, Das Wikipedia-Universum. Wie das Internet unsere Wissenskultur verändert, transcript, 2010. http://www.infoclio.ch/sites/default/files/standard_page/working_paper_geschichte_web2.0.pdf [June 21, 2011]

recensio.net is based upon a two-pillar concept: first, editorial offices of historical journals can publish their [review sections](#) on the platform – either exclusively online, or as part of a ‘hybrid’ format, meaning pre- or post-print. Their Corporate Identity is preserved, as the classical reviews are usually provided in the form of the original PDF files. Secondly, authors are able to present the core statements of a monograph or an article on recensio.net. Gradually, moderated user comments create ‘live reviews’ based on these [presentations](#). With this Web 2.0 idea, the operators of recensio.net are testing a new way of reviewing academic literature in the field of humanities. This offers advantages: the commentator is able to comment on a specific aspect of a book and doesn’t have to write an entire traditional review. This accommodates the needs of today’s time pressed scholars and alleviates the problems traditional journals that find it increasingly hard to find reviewers. Furthermore, the principle of collaborative, flexible and particulate work fits the habits of the younger generation of scholars.

The possibility of republishing their reviews in an international environment, with the effect of them being visible and searchable in full text, is of great interest in the editorial offices of journals and scholarly institutes. Which is why the question of content selection was one of the project’s most crucial points during the launch period. Our main selection criterion was the scholarly quality of each participating journal. Thus it is only possible to stay attractive to the traditional, acclaimed journals publishing book reviews in terms of a future cooperation.

5. Stick To The Principles

There is another reason for having very clear substantial limits when starting a new scientific platform or journal, and for preferring less, but carefully chosen content: amongst scholars, especially in the humanities, there is still prejudice about the Internet and the way it shuffles together a great deal of content which is either not good enough, or not edited well enough to be published on paper. In the case of recensio.net, the limits are defined in the subtitle ‘review platform for European History’: recensio.net doesn’t accept review sections of interdisciplinary journals, those of neighbouring disciplines, or ones that focus too much on regional history. The latter is of course a part of ‘European History’, but especially in the beginning we strictly wanted to hold the course implied in the subtitle, thus preferring transnational views in the concepts of participating journals and their reviews. That way the platform can stay as attractive as possible for future partners from all over Europe.

6. Feasibility

Another reason is a more practical one, which should nevertheless not be ignored. A clearly outlined content selection policy (which limits the content) makes it easier to deal with the technical challenges of building a new platform – especially when the initiators have set themselves high targets concerning full-text search, the enrichment with metadata, or (in the case of recensio.net) the linking of the provided reviews to the respective book title in the library catalogue. For all these reasons, and many more, the issue of time and manpower is a crucial one.

2 Digitisation

Turning to the subject of digitisation and the project [perspectivia.net](#), which is a good example of an online publication platform for the institutes of the Foundation of German Humanities Institutes Abroad (DGIA) and their respective cooperation partners. The aim of the platform is to offer free access to selected research findings of the institutes, be it in the form of monographs, journals, book series, reviews or conference proceedings. In 2008 and 2009, it was funded by the [German Federal Ministry for Education and Research](#) as a pilot project for digitisation processes within the humanities. Several book series and journals were completely digitised for publication on perspectivia.net, and are now searchable in full-text. Other text genres are published in a ‘born digital’ way, namely as html-texts. As a cooperation partner, the Bavarian State Library carries out the digitisation processes as well as the enrichment of the digitised texts with metadata.

The term ‘sources’ within the EUscreen project supposedly refers to audiovisual material from the 20th century. The respective counterpart in perspectivia.net concerning copyright questions would be digitised texts written by historians during the 20th century (including graphical material such as photographs, paintings, maps and so on). This does not necessarily correspond to the classical historical term ‘sources’, meaning primary sources, but in the following, I will use the term in the sense described.

1. Francia

Using the example of the journal 'Francia' I will outline the problems we faced during the digitisation process carried out for perspectivia.net. 'Francia' was first published in 1973 by the German Historical Institute in Paris: it publishes numerous articles and reviews each year, divided into three sections: Middle Ages, Early Modern Period, Contemporary History. In total, about 70 volumes, including more than 1800 authors, had to be digitised before perspectivia.net went online in 2008.

2. Copyright: Text...

The first problem was that of author rights and copyright issues.³ On the advice of an associated media lawyer, each author was told about the proposed plans and was asked for his or her agreement. A time-consuming address research had to be carried out. If an author didn't reply to the enquiry within eight weeks, we implicitly assumed that he or she agreed to the digitisation of the text affected. In total there was only one author out of 1800 who disagreed. Even quite traditional historians know (or have experienced) that the online availability of a text results in its improved perception, given that it is tagged with metadata and that it is searchable in full text.

As is evident, the copyright question can be laborious, especially since more complex points can arise, such as the handling of deceased authors, or those that cannot be traced. Also, we decided to contact authors whose texts were published after 1996 without legally needing to as we felt it important for the relationship between the institute and its authors.

3. ... And Illustrations

Visual material poses slightly more complex problems as in the case of 'Francia' we would have been forced to apply for supplementary licences for every single illustration (e.g. of buildings, of archive material) for each of its 70 volumes. In many cases two rights owners would have been involved – the photographer and the creator of the shown work or document. This subsequent licensing twenty or thirty years after the paper publication was not feasible at all, mainly for two reasons: Firstly, it takes a great deal of time, and secondly perspectivia.net provides the digitised books in Open Access. The reactions of archives or museums are incalculable, often they reject an Open Access publication, or the demanded fees are unaffordable.

4. Licensing

The handling of image copyright on the web is generally at a quite premature, or at least non-uniform, stage especially under the conditions of Open Access. One reason for this is that the national collecting societies often fail to cooperate, even at European level. Consequently, platform operators providing books with visual material have to use immense resources for the supplementary licensing, or they have to work with a makeshift – like we did for the Francia, as numerous right owners outside of Germany were affected: Fortunately the German collecting society (VG-Bild) were sympathetic to our situation, so that we were able to work out an 'all-in-solution', something like a flat-licensing for all illustrations in perspectivia.net-digitised material. Of course our contract does not and cannot cover claims of right owners abroad – but for these cases the German collecting society is willing to help with upcoming conflicts by getting into contact with the affected foreign collecting society (by the way: this has never been the case until today).

There is however one downside to the advantage of not being forced to the labour-intensive supplementary licensing for each single image: The licence fee 'flat rates' are running costs which have to be paid to the German collecting society annually for each book series – beyond the project's funding period. Actually, these costs are not very high, but this point needs to be considered, since such permanent fees have so far not been taken into account.

³ Two exemplary publications dealing with the issue of copyright on the internet (the latter with regard to the term "fair use"): Klaus Graf, *Urheberrechtsfibel – nicht nur für Piraten. Der Text des deutschen Urheberrechtsgesetzes, erklärt und kritisch kommentiert*, Contumax, 2009; Sanfors G. Thatcher, 'Fair Use in Theory and Practice: Reflections on Its History and the Google Case', in Albert N. Greco, eds, *The State of Scholarly Publishing. Challenges and Opportunities*, Transaction, 207-218.

3 Content Genres

When considering an Open Access platform, one needs to think about the question of how suitable certain types of data are for this kind of publishing. Which content genre offers the most 'added value'? How much effort has to be made for placing content online? How well does the target audience accept Open Access publications? For historians, there are four main types of data to be considered: 1. books, 2. e-journals, 3. reviews, 4. video material.

1. Books

Placing printed books online doubtlessly offers added value (namely constant, free availability) but the challenge of clarifying the rights of the publishing houses and authors involved, as well as clarifying potential image right owners, still remains and is often not resolvable, especially in the case of younger books.

2. E-journals

In the case of Open Access e-journals in the historical sciences, we can clearly detect significant added value: evident advantages are fast publishing, perhaps even a more flexible publishing (which allows for updated text versions), the possibility to link the source material, audio and visual material, as well as higher citation rates compared to a paper version. Difficulties remain concerning the impact factor: more so than in the natural sciences, the already mentioned preconceptions regarding the lower quality of texts originally published in an electronic format are – at least in Germany – still existent.

3. Reviews

Another content genre worth mentioning is the field of historical reviewing. Over the last years, historical reviews have established themselves as a genre that is highly suitable for online publishing. The advantage of reviews accessible on the web is obvious: All parties involved have an interest in their free availability. The publishers often see them as a form of good 'advertising' for a history journal, and for the authors of both the reviewed book and the review itself, they create a higher visibility. When the online publication goes hand in hand with a faster publication, the advantage for the user is also apparent, apart from the facilitated access and searchability. Furthermore, the image copyright issue is almost not existent. recensio.net is based on experiences gained on a national level with the review journal *sehепunkte* and goes beyond these, since recensio.net (as mentioned) doesn't publish its 'own (classic) reviews' like *sehепunkte*, but – apart from the idea of providing 'live reviews' – serves as a gateway for existing review journals by bringing together their reviews and making them fully searchable.

4. Videos

Finally, another relevant content genre is video material. The Internet is probably not only a perfect but also the only imaginable place for the publication of video and audio source material for scholars. Nevertheless, systematic access to visual source material has so far been a desideratum, especially from the point of view of contemporary historians.

One point six billion people actively use the World Wide Web around the world [...]. As editors, our goal has always been to preserve and disseminate important historical and literary documents, using the most appropriate tools available. For most of our lives, that tool was the book, and in some cases microform. It simply is no longer the case in 2009. We count ourselves fortunate to sell 1,000 copies of one of our volumes, but over 20 million people have seen just one 1987 video clip of pop star Rick Astley singing 'Never Gonna Give You Up' on YouTube.⁴

4 Contextualisation

Finally contextualisation should be considered, especially with regard to the question of findability of digitised historical content on the web. From the view of an academic platform operator the added value of electronic publishing lies in two main areas: first, the possibility of linking other content (texts, other video material, source material, bibliographies etc.), and second, the chance that content can 'accidentally' be found by the target

⁴ Cathy Moran Hajo, 'Scholarly Editing in a Web 2.0 World', *Documentary Editing*, 31, 2010, 92-103. http://aphdigital.org/people/cathy-moran-hajo/scholarly-editing-in-a-web-2-0-world/#_ftn1 [June 20, 2011].

group. Those unintended discoveries are probably the most important reason for the already mentioned higher citation rates. The main target is, therefore, to 'pick up' the target audience, in our case: to pick up the historian during his 'everyday walks' online.

1. Search Instruments

It is, therefore, particularly important to define the major search instruments of the scientific target audience. Concerning the 'traditional historian', this is still the library catalogue. From the view of recensio.net, it is absolutely essential to link each book review to the respective book entry in the OPAC, and of course to tag each review available on our platform with the relevant OPAC metadata (concerning the classification of time, subject and place for the content of each reviewed book). Within the humanities, this is probably a highly effective way of reaching the scholar's regular attention.

2. Being Linked

To be well linked as a source material platform is one of the most important preconditions for successful contextualisation. Apart from the relative search engines mentioned, this includes the communicative forums most frequently used, blogs and institute websites of the universities, everything that – apart from the library catalogue – belongs to what was described above as 'everyday walks' on the web. These links are an instrument to attract the attention of the potential audience, and at the same time, they are the consequence of that attention. An important factor of this interaction today is social networking: to be seen (as a project) on Facebook or via Twitter feed, which creates visibility amongst the content contributors of tomorrow. In the case of recensio.net, we are even reaching today's contributors, as particularly young scholars commentate on work presentations on the platform, and thus generate 'live reviews'.

3. Cooperate!

The most important way of identifying those 'everyday walks' is the will to allow for cooperation and networking with other well-established, existing, forums, platforms and service. Enabling networking is the Internet's core competence: each partner involved can only win, even though in reality it occasionally takes some time until this awareness drives away the competitive notion. A good concept and strong partners are needed, with libraries playing an increasingly important role regarding e-publishing.

4. PR

The question of findability and contextualisation is closely connected to the question of public relations, especially in the case of international projects. When starting a journal or launching an international platform, one simply doesn't know the individual factors that are essential for the success of the project in Spain, Russia or elsewhere. Thus, it makes sense to invest a substantial amount of work into targeted information management in the specific circle of scholars concerned by the platform's content.

On recensio.net each staff member carries out a country specific web search, which means collecting as many addresses and current publications of historians in different European countries as possible. We then contact these historians, not only to inform them about the existence of the platform, but also in order to tell them that they are personally invited to present their latest work, and of course to comment on the work of others.

At the same time, these country searches result in the gathering of contact details of research institutions, scholarly publishing houses, archives, university repositories and others. The details of these institutions are systematically put together, in order to contact and inform them about the platform, at the same time requesting the establishment of some kind of permanent information about recensio.net for their clients or students. This could be a link on the website, but also by consistently handing out leaflets to graduates or at least at an appropriate point in their 'workflow'. Consequently, a significant part of 'contextualisation' is nothing more than PR.

5. Web 2.0

The question of findability and contextualisation is also closely connected to the Web 2.0 idea: Today an e-journal or a web-platform cannot realistically be launched without giving the users the ability to comment on what they are seeing or reading. Before long it will be completely natural (even for a historian) to comment in an academic way, until there is no doubt that the 'genre' of web-comment does not detract from the credibility, quality or seriousness of the statement it contains. At the moment, the prejudice still exists, especially in the humanities, and especially amongst scholars dealing with 'classical' historical topics. recensio.net does not claim to be a standard solution – but hopes to build confidence by including well-chosen content concerning the 'classical' reviews, and by monitoring and moderating comments on the presentation.

6. Commentators

When it comes to contextualisation, there is nothing more valuable than a well-established system of commentators, who link different places on the web dealing with the same topic. For example, commentators will link the content of a recensio.net review to a video on the same topic published in EUscreen, especially given that the scientific community is of an appropriate size and knows about both platforms in general. The question has to be asked, however: To what extent are there differences between “video-sources” on EUscreen and the ‘illustrated texts’ published on perspectivia.net, or the reviews on recensio.net? What are the differences concerning the contextualisation requirements? The most obvious argument is surely that a video cannot be assigned to a reference point within the most common search engine of a historian, namely the online library catalogue (which is an advantage for both digitised monographs and book reviews, which in contrast can be linked there).

7. Differences

As a conclusion it is necessary to again take up the point of citation rates, which is an important argument for e-publishing of texts and an important factor concerning their contextualisation. Where are the differences in ‘being cited’ between a text and a video? And – moving one step back: Before a text can be cited it has to be found. That is why the accidental finding via full text searchability and search engines is very important for the visibility of a text on the web. How can the ‘strategic disadvantage’ of video material (i.e. its lack of full-text searchability) be compensated? And finally: Are there existing technical solutions for a kind of audio- or video-OCR which generates text out of audio tracks and adds this text to the metadata of a video?

There will be no panacea for the construction of sustainable scholarly platforms. Each discipline has its very own specific needs in terms of its structure and network strategy. In Germany, some of the very early tools for historians, such as the journal [sehpunkte](#) or the information service [H-Soz-u-Kult](#), have not only managed to survive, but have even institutionalized themselves into hubs which by now hardly any scholar can ignore. This article has tried to describe four essential issues which should be considered during the conception of new services. Solution strategies can only be individual, but above all should remain flexible. Just as the web itself is subject to constant change and renewal, habits regarding the consumption of content, as well as reading and reception habits, change.

Biography

Lilian Landes studied art history and archaeology at the universities of Marburg and Munich. In Munich she received the doctor’s degree with a dissertation about social genre painting in Germany in the 1840s. She joined the Centre for Electronic Publishing at the [Bavarian State Library](#) in 2008 after having worked four years at the [University of Munich](#) (Collaborative Research Centre “Early Modern Period”). At the BSB she has been a member of the academic editorial team of [perspectivia.net](#). Since 2010 she has been co-ordinating the [recensio.net](#) project (review platform for European History).