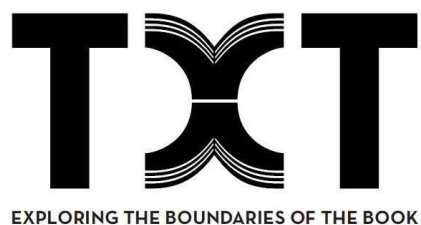


Cover Page



The handle <http://hdl.handle.net/1887/28849> holds the collection of TXT in the Leiden University Repository.

This document has been released under the following Creative Commons license



Defragmenting Digitised Manuscripts Sources

The DMMmaps: A Unified Portal to Medieval Manuscripts

By Giulio Menna and Marjolein de Vos

Managers and Editors, *Sexy Codicology*

Introduction

A search with the terms ‘Digitised Medieval Manuscripts’ in search engines leads to well established libraries on the first results page: The British Library, The Houghton Library, et al.¹ in combination with other results, such as articles and blogs on the same matter. These search results make finding other online repositories difficult due to the fact that many digital libraries have not been optimised for easy access by search engine crawlers. Sanderson et al. have estimated that only around 1% of existing medieval documents have been digitised to date, and most of the effort has been spent in taking care of the visual presentation of the digitised projects. Many institutions, furthermore, have used independent approaches to the subject.² These various factors have led to a fragmentation of digitised resources over the internet, which makes discovering and browsing medieval manuscripts difficult.

To battle this issue the DMMmaps Project was created. The DMMmaps is an independent project that aims to collect web-links to digitised medieval manuscripts repositories and make them available on an interactive map. The project originated as an attempt to unify well-known digitised sources already available (*E-Codices* in Switzerland, *Digital Scriptorium* in the United States and *La Bibliothèque Virtuelle des Manuscrits Médiévaux* in France) along with many other less known repositories and present them in a single place, in order to create a faster and easier way to find and access these sources. The DMMmaps are aimed both at scholars and at enthusiasts interested in medieval manu-

scripts. It is designed as a free service with further potential developments and contributions from anybody interested.

This paper will explain the various aspects of the project: how the DMMmaps originated and were initially developed, the reasons behind the choice of a virtual map as an interface, the way that DMMmaps operates, difficulties that have been encountered, and the crowdsourcing aspect. The importance of a social media presence will be also discussed throughout the paper.

How the DMMmaps project originated and blue the use of maps

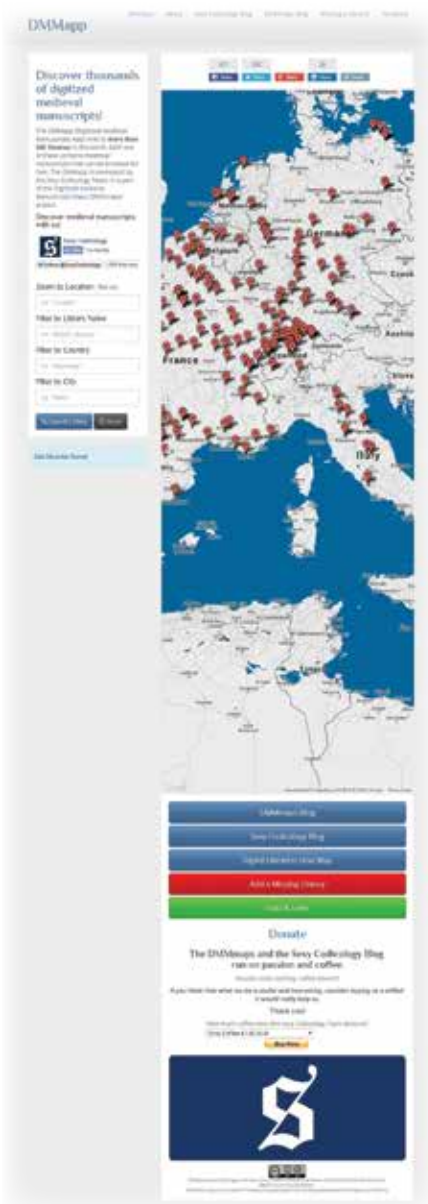
The DMMmaps is the direct result of another project managed by the authors of this article: The Sexy Codicology Blog. This blog aims to promote an interest in medieval manuscripts and relies heavily on the use of digitised manuscripts images to create articles and videos. While searching for content to post, digitised repositories that were arbitrarily discovered were added to a list of 'Favorites' in the browser. This led to the creation of a long list of links that was difficult to browse and not shareable with other interested parties. The easiest solution to this issue would have been to create a simple database that would have allowed users to filter digitised repositories based on nation, location and name.

What this solution lacked was interactivity and navigability: the user would have been presented with either a long list of names, links and information, plus the filtering system; or the filtering system without the information of all the libraries available. With these issues in mind, a small test in which these librar-

ies would be represented on a digital map was created, and the basic information concerning the digital libraries (i.e. name of the library, city, nation, and number of digitised manuscripts) was added. The service used at this time was Google Maps Engine Lite.³ This 'alpha version' of the DMMmaps project underwent private testing for two weeks. The solution appeared to work. The user was not overwhelmed by a huge amount of information from a large number of libraries, the navigation of the map was easy and straightforward, and the service was free to use. After this trial period, an early example of the project was made public. The response of scholars and enthusiasts were key in order to understand if what was being done was useful or not.

From July 7th 2013 until August 7th 2013, the publicly available maps were viewed 838 times. This became 1885 views in the period from September 8th 2013 until October 6th of the same year. By December 20th 2013 the DMMmaps project was visited a total of 5116 times. On this date, the project, which was hosted on the Sexy Codicology website, was moved to its current domain: digitisedmedievalmanuscripts.org

The response of the public was exceedingly positive but as a consequence, the limits of the Google Maps Engine were reached. The number of libraries that were added forced the creation of four different maps for the libraries in Europe alone, causing confusion to the visitors who would not see a library on a map and would send a link for it to be added. Because Google Maps Engine would only allow a hundred pins per map, not all libraries of the world could be displayed on one map. Even one map of, for example



Europe, was not possible. This forced the complete redesign the DMMmaps. It was now necessary to develop a customised app that would be easier to use, aesthetically appealing, and also usable from a mobile device (smartphone or tablet), which could also accommodate more than 300 libraries on a single page.

Google maps API was chosen in order to modify the look and feel of the map in

JavaScript. The database that had been previously created with the links, number of digitised manuscripts, and geographical position of the libraries was transferred into Google Fusion Tables. Together, the Fusion Tables and the API visualise the information as a map with pins. The app was then uploaded to *digitisedmedievalmanuscripts.org*, along with a WordPress installation for blogging purposes, followed by search engine optimisation and made available to everyone. The total cost of this operation was 18 euros, the price for renting space on a shared server for a year.

Collecting the links to digitised medieval manuscripts and the social media factor

The collection of links to digitised books is mainly done through crowdsourcing. These contributions, containing links and basic information about digital repositories, are sent to the DMMmaps project in a variety of ways, primarily with the use of a simple web-form on the project's website. In this form the contributor must add the name of the library that he would like to see added, the nation and city where the institution is located and the link to the digitised manuscripts. There is also an option for users to add to the number of digital objects available and post further remarks. Any user aware of a missing digitised manuscript collection can contribute to making the DMMmaps complete. This method led to 67 different contacts and the addition of around 50 libraries to the DMMmaps. Another way in which the links to digitised manuscripts were acquired is by the active research by the project's authors. This was espe-

cially true in the beginning of the project due to the fact that there were not enough users to feed the links. Many tips were also received through social media interactions. ‘Tweets’ led to adding four libraries to the database; three were added through the Sexy Codicology Facebook Page.

This ‘social media’ aspect of the project is essential and cannot be underestimated: it is only thanks to the power of social media that the DMMmaps project has reached more than 30,000 visitors in little more than four months. Statistics coming from the digitisedmedievalmanuscripts.org website show that, in the time period from the 1st of January 2014 until April the 30th 2014, 66.6% of the visitors of the website came via the Sexy Codicology social media channels. Organic Search (i.e. referrals from Google, Bing, Yahoo and other search engines) accounted for only 4%; referrals from other websites were 5.9%; direct visits (i.e. visitors who bookmarked the website) accounted for a total of 22.9%.

Another important aspect of the DMMmaps project is that it is backed by the Sexy Codicology blog. This is a project that we created one year before launching digitisedmedievalmanuscripts.org. Sexy Codicology is a blog concerning medieval manuscripts. With over 4,000 likes and



1,500 followers on Twitter, it helped generate public awareness of the DMMmaps project.

Thanks to the careful social media strategy and crowdsourcing, today the DMMmaps database connects to 312 libraries, containing at least 30,000 digitised medieval manuscripts.⁴ The project has received at least 57 crowd-sourced links to digitised repositories and, furthermore, institutions such as the Reynolds Library in Manchester, have kindly requested to be added to the database to enhance the visibility of their collections.

Now that the links’ acquisition has been explained, it is time to explain how they are entered into the database. The process is straightforward and efficient: the information sent by a user is inserted in a datasheet (Google Spreadsheet file), latitude and longitude are added with Google Earth. The same data is then inserted into the Google Fusion Tables file, through which the link is ‘pinned’ to a

virtual map that can be navigated by the user. Clicking these pins shows the following information of the repository:

- *Name and location of the institution that hosts the digitised manuscripts.*
- *Number of digitised objects available on the institution's website (including date of availability).*
- *Link to the website of the digital repository.*

Copyright license & budget

Although the project is maintained privately, the database on which the project is based is shared under a Creative Commons Attribution - Sharealike License, allowing free use of the data acquired.⁵ The intent of this approach is to push as many people as possible to create their own project, based on the data we have created. The choice of using a free culture license came naturally due to the fact that the project is based on a series of web links, which are presented in an effective and easy to use format. It would be possible for anyone to locate these websites, copy and paste the links into a new database and do the work all over again. It is the hope of the creators of this project that someone, one day, will step up and improve what has been done until now, maybe creating an even better version of the DMMmaps.

The DMMmaps currently receives no funding and the project is made possible by using free software available on the Internet. Specifically, the project uses the following software and services:

- *Google Spreadsheets*
- *Google Fusion Tables*
- *Google Maps API*

The interface of the app is simple HTML5

and CSS3, based on a freely available template by Derek Eder.⁶ Although free to use, the products offered by Google guarantee the flexibility, power and possibilities given by paid software. These three services are also characterised by a distinctive ease of use; the DMMmaps were, in fact, created in a matter of days. There are, however, still issues with using these programs. Google has recently shut down a series of products of its own family (including Google Reader, Google Latitude, iGoogle in 2013) and there is no guarantee that the products on which the DMMmaps are based will be active and available indefinitely. This issue could be fixed by creating a database that does not rely on Google services, but on well-established and standardised programmes (or by hard coding the data directly into the script). In the specific case of DMMmaps this was not possible, due to the fact that including all the collected data in the script directly would make it extremely slow to load. Using a professional database was also impossible due to the lack of funds available.

Issues

An ambitious project like DMMmaps is always met with difficulties. Currently the major issues that have been identified are as follows:

Generic issues:

It is time consuming to manage social media channels.

The maintenance of active and interesting social media channels is an inevitably time-consuming task. The time involved in searching, sharing, and resharing content is only partially mitigated by the use of scheduling tools.

The need for specific technical knowledge to manage the website and integrate the results.

Although the DMM-maps project is aimed at a purely humanities subject, the development is based on the capabilities of its developers to use web services, coding, and scripting appropriately to the needs of the digital project.

This leads to the consideration that as cultural heritage becomes increasing digital, it is becoming more essential for humanities scholars to code and create basic applications. Improving digital literacy is crucial for humanities scholars in order to allow more effective sharing of research and studies.

Gaining momentum and audience attention.

As mentioned earlier, the Sexy Codicology blog is a powerful ally to the DMMmaps project due to its consistent number of followers that are specifically interested in medieval manuscripts and ancient books. One can easily understand that if the project did not originate from such a favorable position, it would have had to work much harder to gain the same interest and following. This thought leads us to several considerations on the importance of social-networking for digital humanities projects.



Project-specific issues:

Difficulty in counting the total number of digitised objects.

A core objective of the DMMmaps is to collect data in order to better understand the current situation of manuscript digitization in the world. Part of this data collected therefore concerns the number of digitised manuscripts that can be accessed in any given library that is linked on the map. What we often observe is that few institutions clearly indicated the number of digitised manuscripts they have available online. This noticeable lack of data concerning online digitised manuscripts prevents researchers from creating interesting statistics and correlations. For example the Heat Map created for this project, clearly indicates which countries have the most institutions involved in the digitization process. What would have been interesting to create is another heat map showing which parts of the world

hold the highest concentrations of digitised medieval books, and then correlating this with the aforementioned map of digitised institutions.

Broken links to digital repositories (absence of permalinks).

Although not a direct fault of the DMMmaps project, broken links are a major issue. To give an example, recently the Vatican Library migrated all its digitised medieval manuscripts to a new website, without giving prior notice or, more importantly, without implementing a 301 redirect from the old page. This meant that users will no longer be able to access digitised Vatican manuscripts through their bookmarked webpage. This also resulted in a blank page being linked from DMMmaps. Once the situation was noticed, the link was updated in a matter of minutes.

It is possible to quickly report broken links through the app, but the only true solution would be the use of proper code and solutions on behalf of the webmasters and web developers of the digitised libraries.

Conclusion

First and foremost, the experience with the DMMmaps project shows that a useful digital humanities resource, usable by researchers and enthusiasts, does not necessarily have to be expensive and complex. As we mentioned earlier, the cost of developing and maintaining the project website is merely 18 euros per year. By taking advantage of free programs readily available on the internet, it is possible to develop a professional project without necessarily possessing a high level of expertise in coding and developing app. The risk in

using free software is that the infrastructure on which a project is developed might not be available for long. Nonetheless, this remains a highly viable way for young researchers and independent scholars to become noticed in the professional world.

The development of the DMMmaps also highlighted the highly fragmented world of digital libraries. Often, the effort made to digitise medieval manuscripts and make them publicly available, is hindered by poor web design and deficient search engine optimization. In fact, very few of the manuscripts that were added to the project were discovered via search engines. Most of the links that were added to DMMmaps came from crowd sourcing. More specifically, the researchers and librarians working for institutions that housed digitised manuscripts have been the most active contributors. It is still too early to make any conclusive claims about the effectiveness of the strategies we have discussed in relation to the DMMmaps project.

However, the number of crowd-sourced contributions we have received appears to indicate that an approach based on a strong social media presence is an effective way of creating a complete database of digitised repositories. Social media allows potential contributors to be in direct contact with the project and share their suggestions for future improvements.

This DMMmaps project will continue to be developed and updated in order to become an even more comprehensive tool for both the scholar and enthusiast, in search for digitised medieval manuscripts. ■

Notes

1. The results taken in consideration here are those of a search done in Google's "Incognito Mode" in order not to have personalised results. For further information on personalised results: <https://support.google.com/accounts/answer/54041?hl=en>
2. Sanderson, Robert et al. 'Sharedcanvas: a collaborative model for medieval manuscript layout dissemination.' *Proceedings of the 11th annual international ACM/IEEE joint conference on Digital libraries* 13 June 2011, 175-184.
3. This service can be accessed here: <https://mapsengine.google.com>
4. A precise count of the number of digitised objects is currently impossible to obtain. Some libraries count the number of digitised folia; others do not insert the exact numbering of digitised objects at all. 30,000 is the sum of the digitised books from libraries, which have added a precise counting of digitised manuscripts.
5. Creative Commons. 'About the licenses.' *Creative Commons* (2009).
6. This useful project can be found here: http://derekeder.com/searchable_map_template/