**TERRAS — Collaborative**

<3 figures>

**Collaborative Digitisation: UCL’s Medieval Manuscript Fragment Project**

Terras, M.; Graham-Matheson, H.; Furlong, G.; Wright, S.; Makin, K. and Gibson, A.

Many collections of medieval manuscripts are now being digitised and showcased online, but in order to do so, a complete pipeline of identification, cataloguing, imaging, and delivery is needed, which can be prohibitively expensive and time consuming. Teams seldom have all the skill sets in place to deal with the identification of complex historical texts. It is also rare to attempt any advanced imaging techniques to aid the reading of manuscript material. We report here on an interdisciplinary collaboration to analyse fragments of unknown medieval manuscript material held in UCL Special Collections. We demonstrate that a centre for digital humanities can provide the catalyst to undertake such a small, yet intensive, digitization project within an institutional context, and, working in joint collaboration, can aid in the process of digitization, communicate with a wider online audience to help with identification of material, and promote interdisciplinary working to prepare material for further research in heritage science.

**The Collection**

UCL Special Collections,1 one of the foremost university collections of manuscripts, archives, and rare books in the United Kingdom, hosts fine collections of medieval manuscripts, early printed books, and archival material. A collection of 157 medieval manuscript fragments have been in the possession of the Library for decades, but have never been fully described and are generally unknown to the wider community. They were purchased in the early 20th century in an initiative led by a professor of German studies at UCL, Robert Priebsch, who had long been trying to secure a group of manuscripts for palaeographic study (Munby, 1960, 38). Money was raised by subscription amongst UCL staff (Munby, 1960), and it is thought that this collection was purchased at auction in Bonn in 1921; unfortunately relevant records did not survive the war (Munby, 1960, 42). Although mentioned briefly in a catalogue of UCL’s holdings in the 1930s (University College, London and Coveney, 1935), the manuscripts are yet to be fully described and researched. Mostly originating from Germany, they have been found to contain rare examples of early musical notation as well as legal, religious, medical, and other texts from a range of dates (mainly 10th–14th centuries), styles, and different languages (including Latin, German, Hebrew, English, French, and Greek). Consistent with having been used in bindings for later printed works, the majority of the fragments have significant damage, most usually remnants of, or loss of text removed by, adhesive. They are variously cut, torn, and faded, with pest damage and various accretions.

**Digital Humanities and Digitisation**

Although most effort in the digital humanities is focussed on the production, analysis, and visualization of text, there is a growing interest in the community towards digital imaging. Digitisation can produce adequate scholarly surrogates of historical documents (Terras, 2008). However, improvements in digital image processing and analysis have allowed the development of a number of techniques that can reveal a greater wealth of information about the originals, beyond traditional digitisation technologies (MacDonald, 2006). Scholars have been able to image, analyse, and recover more information from historical texts, primarily using a technique called multispectral imaging (Chabries et al.,2003; Salerno et al., 2007; Tanner and Bearman, 2009). Over the past few years, UCL has been building up its expertise in cultural and heritage imaging, working on a range of projects that involve developing and analyzing capture methods for primary historical texts (for example, see Pal et al., 2014; MacDonald et al., 2013; Giacometti et al., forthcoming).

We conceived the Fragment Project as the inaugural scholarly programme of the recently opened UCL Multi-Modal digitisation suite,2 a facility for teaching and research in digitisation technologies jointly supported by UCL Faculty of Arts and Humanities,3 UCL Faculty of Engineering Sciences,4 and UCL Library Services5 coordinated by UCL Centre for Digital Humanities,6 completed in September 2013. In working together on the digitisation of the manuscript fragments, we aimed to make the fragments available to a wider scholarly audience for the first time, help in their identification and classification, but also to categorize the physical properties of this set of primary source material held locally in UCL Special Collections, so that we can use the fragments in our research aimed at developing new imaging techniques for recovering text in deteriorated historical manuscripts.

**The UCL Medieval Manuscript Fragment Project**

This project was given a small amount of seed funding from both the Library and UCL Centre for Digital Humanities (totaling £3000) to allow the classification, identification, digitization, and cataloguing of the collection. The first phase of the programme, undertaken between November 2013 and April 2014, involved surveying all fragments in the collection, to identify—where possible—the language, content, and date of the fragments, and to quantify the physical condition of each (for example, Cracks, Flaking, Fading, Abrasion, Ink Water Solubility, Loss, Ink Offset, etc.). We were lucky to employ, part time, a recently finished PhD student with wide-ranging linguistic and palaeographic expertise to undertake this overview. Fifty detailed descriptions of selected fragments were produced and used as the basis for a full archive catalogue, which was produced in accordance with international standards using CALM software.7 Two hundred seventy-nine high-resolution images of the collection were produced.



Figure 1. An example of a fragment identified and digitized. Thirteenth-century leaf from a hymnal, parchment. Text in a gothic script in two columns interspersed with musical notation in a German gothic ‘hufnagel’ style on a 4-line stave with C marked. UCL Library Services Special Collections MS.

Although we were successful in identifying the source of most of the fragments (which are varied: a fragment of St Paul’s Second Epistle to the Romans, a miniscule portion of the Greek play *Medea* from the 4th or 5th century AD, a Gregorian chant based on Psalm 65), the digital humanists in our group also encouraged the use of social media to connect with a wider community of scholars (UCL does not have a music department, and we had no expertise in our team to help with identification of the music fragments). Others have had success in the identification of medieval fragments via crowdsourcing (Erwin, 2013), so we turned to the hive mind. Liaising with groups such as the Plainsong and Medieval Music Society on Facebook (https://www.facebook.com/groups/108654282490212/) and Musicologie Médiévale (http://gregorian-chant.ning.com/) allowed us to communicate with other manuscript experts and gain further insights, identification, and classification.



Figure 2. Successful outreach via social media to aid in further identification of

fragments.

Many of our fragments are now online and can be seen at bit.ly/uclmedievalfragments. This phase of the programme will have several additional benefits for teaching, research, and public engagement. The manuscripts will be showcased in exhibitions and disseminated online, forming a unique set of teaching and research materials. However, the innovative approach of this activity is in being able to prioritise which manuscripts will be most useful to focus efforts on, in order to help read damaged and deteriorated texts via advanced multi-modal imaging techniques, and at the time of writing we are moving into the final phase of the project, quantifying best practice in using multispectral imaging on damaged and abraded text. We now have a range of locally held, categorized medieval fragments with which to work as we perfect our techniques before visiting other archives to image historically important manuscripts.



Figure 3. MS FRAG/MUSIC/15, an antiphonal featuring Gregorian chants. Text and square musical notation in black with 4-line staves ruled in red. Only 40% of text legible: an example of a fragment we can work on with multispectral imaging to improve legibility of the text.

**Conclusion**

Framing the project as a digital humanities one opened up access to financial and physical resources provided by UCL Centre for Digital Humanities, gave easy access to institutional infrastructure for setting up and maintaining digital projects (such as data backup for digitisation), allowed identification and employment of staff to aid in the surveying and identification of the fragments, aided in reaching out to a wider online community to help identify fragments of unknown provenance, and allowed a large part of the digitization of the collection to be carried out. In doing this work we are now prepared to undertake novel imaging research with the collection (demonstrating the groundwork that is necessary before undertaking advanced cultural heritage imaging). It should be stressed that we could not have undertaken this project without the full support of the Library, whose own investment allowed refining of the collected data, cataloguing, mounting of the digitized material into the institutional content management system, and further digitization. By pooling resources and expertise, we have undertaken identification, cataloguing, and digitization of a collection, for a mere few thousand pounds, whilst preparing for a more advanced phase of imaging research. The project demonstrates, then, that centres for digital humanities, working in equal collaboration with institutional partners, can provide complementary resources and expertise to support complex digitization processes, whilst this mutual relationship can lead to future novel research endeavors.

**Notes**

1. http://www.ucl.ac.uk/library/special-collections.

2. http://www.ucl.ac.uk/dh/facilities/digitisation-suite.

3. http://www.ucl.ac.uk/ah.

4. http://www.engineering.ucl.ac.uk/.

5. http://www.ucl.ac.uk/library/.

6. http://www.ucl.ac.uk/dh.

7. http://www.axiell.co.uk/calm.

**References**

**Chabries, M., Booras, S. W. and Bearman, G. H.** (2003). Imaging the Past: Recent Applications of Multispectral Imaging Technology to Deciphering Manuscripts. *Antiquity,* **77**(296): 359–72.

**Erwin, M.** (2013). Crowdsourcing the Arcane: Utilizing Flickr (and Google) to Describe Medieval Manuscript Fragments. Updated transcript of talk given at *Society of Southwest Archivists Annual Meeting*, Austin, TX, http://micahcapstone.wordpress.com/2013/06/01/crowdsourcing-the-arcane-utilizingflickr-and-google-to-describe-medieval-manuscript-fragments/.

**Giacometti, A., Terras, M. and Gibson, A.** (forthcoming). Objectively Evaluating Text Recovery Methodologies for Multispectral Images of Palimpsests. *International Journal of Heritage in the Digital Era*, 15th issue dedicated to Computer Vision in Cultural Heritage.

**MacDonald, L. W.** (2006). Digital Heritage: Applying Digital Imaging to Cultural Heritage. Elsevier, Amsterdam.

**MacDonald, L., Giacometti, A., Campagnolo, A., Robson, S., Weyrich, T., Terras, M. and Gibson, A.** (2013). Multispectral Imaging of Degraded Parchment. *Computational Color Imaging, Lecture Notes in Computer Science,* **7786**: 143–57.

**Munby, A. N. L.** (1960). *The Dispersal of the Phillips Library.* University Press, Cambridge, http://books.google.com/books?id=91e5AAAAIAAJ.

**Pal, K., Schuller, C., Panozzo, D., Sorkine-Hornung, O. and Weyrich, T.** (2014). Content-Aware Surface Parameterization for Interactive Restoration of Historical Documents. *Computer Graphics Forum (Proc. Eurographics),* **33**(2).

**Salerno, E., Tonazzini, A. and Bedini, L.** (2007). Digital Image Analysis to Enhance Underwritten Text in the Archimedes Palimpsest. *IJDAR* **9**(2–4): 79–87.

**Tanner, S. and Bearman, G.** (2009). Digitizing the Dead Sea Scrolls. *IS&T Archiving 2009, Final Program and Proceedings*, pp. 119–23.

**Terras, M.** (2008). *Digital Images for the Information Professional*. Ashgate, London.

**University College, London and Coveney, D. K.** (1935). *A Descriptive Catalogue of Manuscripts in the Library of University College, London*. Printed for University of London, University College.