Studying the Book of Hours:
Creating Tools for Scholars and Evaluating Manuscript Digitization

In this project we propose to develop tools to allow the analysis of scanned manuscript pages as images. Our goal is to provide both tools for segmenting, organizing and navigating the images, and tools for recording scholarly notes and observations in both the public manuscript database, and in a private database. In addition to providing tools, this work will allow us to assess whether existing scanned manuscripts have been digitized in a manner that supports this type of scholarly examination. We will use as a case study scanned English books of hours.

More books of hours remain in modern libraries than any other kind of book from late medieval England: almost eight hundred manuscript volumes, and many thousands of printed ones. Their survival rate suggests that they were very widely read; a recent exhibition at the J. Paul Getty Center has described the book of hours as a "medieval bestseller."[1] Because books of hours were the main textual object used by medieval laypeople, so often read by women, and even the basic book of children's learning (the name *primer* perhaps derives from the hour of *Prime*), they were fundamental to the reading practices of a variety of audiences. Yet critics have failed to ask how the popularity of the book of hours shaped the literate culture within which it grew up. Because the constitutive text of the book of hours—the Little Office of the Virgin—is an unvarying set of Latin prayers, modern readers have deemed the volumes verbally uninteresting. But the book of hours comprises more than this text alone: as a compendium of multiple texts and systems of images, the genre depends upon a rich hybridity of representation. These manuscripts, which provided the most common experience of books in the late medieval period, informed literate culture at large with their complex understanding of what it means to read. The book of hours, then, offers exciting opportunities for thinking anew about some of the most important innovations of late medieval literary culture.

Jessica Brantley (English) is currently exploring the importance of the book of hours for literary history. Rather than offering a history of art or a history of prayer—the more common rubrics for approaching these volumes—her study excavates the histories of reading that are manifest in this uniquely large textual archive. She is examining the nearly 800 books of hours extant that were made in England or for the English market: a copious archive that allows for investigation of connections both broad and specific between the culture of reading shaped by these books and vernacular literature in the later Middle Ages. Perhaps surprisingly, English literary culture is visible within the manuscripts of the hours, and traces of hourly prayers and their books also appear within vernacular texts.

This project provides an exciting opportunity for exploring the possibilities for digital analysis of a large and relatively standard group of objects. Because of the great visual interest of books of hours, we will explore methods of quantifying and searching for scribal or artistic "signatures" through visible shapes—both the shapes of letters (matching scribes) and the shapes of drawings (matching artists). More important, we will develop a tool for comparative analysis, one that would allow for two image to be compared on the same screen. A space for creating a virtual

---

[1] *The Medieval Bestseller: Illuminated Books of Hours*, an exhibition held at the J. Paul Getty Center, Los Angeles, CA, October 29, 2002-January 26, 2003. See also *Medieval Bestseller: The Book of Hours*, an exhibition held at the Pierpont Morgan Library, New York, NY, September 17, 1997-January 4, 1998.

"collection" of all of the manuscripts relevant to a single project, where they would be subjected to some kind of comparative qualitative analysis, would be even more useful than new digitization.

The goals in studying the images are:
1)   Many illustrations are scattered through the books, often unrelated to the text. Studying how these illustrations are the same or vary within a manuscript or through different manuscripts will allow the scholar to gain insight into the development of these books, and the development of books in general.
2)   Many decorative borders are used in the books, quite standard but with discernible and revealing differences (Scott, 2002). How are these borders the same or different in different manuscripts, and how might establishing similarities allow for more secure dating and localization?
3)   Even though the textual content of books of hours for the most part is the same set of prayers, there are instances of additional texts in the books. The texts included can also provide insight into the more general development of books and their use. For example, we can locate likely locations of verse in the masses of manuscript pages from the layout of forms in the image.

To facilitate this image analysis, we will build on previous work in manuscript analysis (e.g. Le Bourgeois and Kaileh, 2004) to develop custom algorithms to segment the manuscript images into areas of text and areas of individual illustration. We will develop and test algorithms for measuring similarity of illustrations based on interaction with the scholar and their assessment of what is interesting in comparing two illustrations. We will test various detectors of text layout to see if we can indeed accelerate the location of poetry within the books.

Comparing large numbers of manuscripts, forming questions, and making observations is a large record keeping project. Therefore, an important aim of this project is to develop a way to organize all of a scholar's notes and provide links to the manuscript(s) related to the notes. In some cases it will be desirable to place annotations in the public data base for the reference of other scholars. For the most part, however, we will focus on creating a private work environment that will enable a scholar to make notes and easily recall and relate them to one another.as a larger work is being prepared. We propose creating an environment based on interaction with the scholar and observation of their working method. On a high level this is following previous work on creating a work environment for classicists (de la Flor et al., 2010).

Our plan of work will be to start with a set of manuscripts that have already been digitized. This will allow us to assess whether the quality and metadata stored in these existing projects was adequate for this scholarly study. If inadequacies are found, we will propose a new digitizing protocol for additional works that meets the need of the scholar.

We would ideally make use of a graduate assistant, to help determine which of the books of hours are already digitized, and to construct a list of what digital resources are already available for this study. {How many hours?}

References

Frank Le Bourgeois and Hala Kaileh, 2004. Automatic Metadata Retrieval from Ancient Manuscripts, in Marinai, Simone and Dengel, Andreas, Eds. Document Analysis Systems VI , Lecture Notes in Computer Science, Vol. 3163, Springer Berlin / Heidelberg, 75- 89

Grace de la Flor, Paul Luff, Marina Jirotka, John Pybus, Ruth Kirkham, and Annamaria Carusi. 2010. The Case of the Disappearing Ox: Seeing Through Digital Images to an Analysis of Ancient Texts. In Proceedings of the 28th international conference on Human factors in computing systems (CHI '10). ACM, New York, NY, USA, 473-482.

Kathleen Scott. *Dated & datable English manuscript borders, c. 1395-1499*. Bibliographical Society, 2002.