[1] Title

[2] Some of the traditional roles played by printed codicological descriptions have not changed in digital environments. Descriptions formalize an approach to and vocabulary for understanding cultural artifacts, and they provide an expert opinion on the origins and status of manuscript books for the benefit of scholars who are unable to consult the original objects and/or non-expert users who lack the necessary skills to make such judgments for themselves. Because they typically summarize dates, origins, owners, and contents of books, descriptions also serve as useful preliminary resources for researchers looking for information that will suggest which volumes, collections, and repositories are most likely to reward further time and effort.

[3]Codicological descriptions are usually characterized by a highly specialized and specific vocabulary—developed and augmented over the years by curators, codicologists, art historians, and others—and terse prose entries that make highly efficient use of space in printed books. Abbreviations and formulae are common, as they facilitate conveying a considerable amount of information in a brief space.

[4]These features are seen clearly in the following excerpt

from a description of M. 948, a Rose manuscript held by the Morgan Library & Museum

[5] Such descriptions have traditionally met (and continue to meet) the needs of two types of users. The first is the visitor to the library who wishes to use the description as a guide to a manuscript being consulted in person. Information such as the name of the scribe responsible for the manuscript, the location of—and text lost as a result of—missing leaves, and the distribution and relative merit of the work of the two artists facilitates and expedites the work of most researchers, and is particularly valuable in enabling the work of those who wish to consult the manuscript for literary, historical, or other reasons, but are not themselves equipped with the specialized knowledge to make such judgments.

The second type of user whose needs are met by such descriptions is the researcher who is studying the manuscript remotely, and thus needs information that would otherwise be available only if the manuscript were at hand. The fact that the manuscript contains two columns of 33 lines each, for example, or that there are "numerous gold initials against alternating red and blue backgrounds throughout" is information that one does not need to provide to a library visitor who has the manuscript in front of her. In such cases, the language of

codicological descriptions has needed to be precise and clear because it needed to convey an image of an original object that a user often could not see in person.

[6] In digital environments, we encounter new forms for both codicological descriptions and the objects they describe. As Daniel Pitti has observed, "[i]n order to apply computer technology to humanities research, it is necessary to represent in machinereadable form the artifacts or objects of primary or evidentiary interest in the research, as well as secondary information used in the description, analysis, and interpretation of the objects" (474).

[7] Three categories of information. The first of these—the dissemination of specialized knowledge—remains relatively unchanged; whether working in a physical or an online library, many users will need the combined paleographical, codicological, literary, and art historical knowledge found in descriptions such as that of M. 948 above. Such information is of course more easily searched, mined, and disseminated in a digital environment, but this is true much more broadly of marked up texts of all types, and thus need not detain us here.

[8] The second category is information that refers to the physical nature of manuscripts, and hence is not available via digital surrogates. This includes physical measurements of bindings, folios, and text blocks, tactile information such as the thickness of paper or whether one is seeing the hair or flesh side of parchment, and a reliable collation of the book, which necessitates physical inspection. This category of information also serves as a check against distortions to our understanding of physical objects that occur in electronic environments. Online libraries and archives are frequently equipped with tools for manipulating images, such as the ability to zoom, pan, and rotate. While such technologies are enormously useful to a researcher wanting to conduct a detailed analysis of a miniature or marginal inscription, they also tend to distort a sense of scale, both within one book and between multiple books. Digital repositories, meanwhile, are subject to mistakes that look remarkably similar to those made centuries ago in scriptoria and binderies. Instead of mistakes in foliation or pagination, files are misnamed. A break in a digital codex might as easily be the result of a lost file as a lost leaf in the physical book it represents. And rather than a binder misordering his gatherings, we might find files sequenced incorrectly. Descriptions made from physical books therefore serve as a means to diagnose and correct such problems.

[9] The third category pertains to information that previously was included to meet the needs of those researchers studying manuscripts remotely through descriptions of them; it is in this category that we witness the most fundamental changes in the purposes and uses of codicological descriptions in digital environments. This category concerns information that, in printed descriptions, was designed to summarize and provide details of the physical appearances of manuscripts. Needless to say, the need for such information is substantially lessened when descriptions are accompanied by digital images; that the text of M. 948 is in two columns or that there are "numerous gold initials against alternating red and blue backgrounds throughout" is now attested by the images themselves, and thus there is not the same need for this information in the description. But this information has gained new usefulness even as it has lost much of its original purpose, for it now serves as a means for sorting, classifying, and comparing collections of manuscripts.

[10] In addition to changes in the purposes and uses of codicological descriptions, their relationship to what they describe has changed in number and complexity. Printed codicological descriptions exhibit a one-to-one relationship to the manuscript books they describe, offering a summary and analysis of the book's physical and textual properties.

This is not meant to imply, of course, that the descriptions themselves or their relationships to the codices they describe are in every case simple.

In digital archives comprising images, however, this one-to-one relationship is supplanted by a one-to-many relationship. At a minimum, the original book, the codicological description, and the images that constitute the surrogate book each present relationships to the other two. In such an environment, the description describes not only the original book, but also the surrogate. While the original codex maintains an ultimate authority in that it possesses the ability to show whether a codicological description and/or the surrogate codex is somehow faulty or incomplete, the reality, and indeed the very goal, of most digital libraries is that far more people will use the digitized description as a guide to the surrogate book than would ever be able to use it as a guide to the original artifact. As such, a description in a digital environment should work equally well as a guide to both.

The original codex and its surrogate images also participate in one-to-many relationships. The images, like the description, are a representation of the artifact; in turn, the images are described by and linked to the description. The original codex, meanwhile, stands in a set of new relationships to virtual versions of itself. But of course, this

model is frequently complicated still further, as when transcriptions and other metadata offer new sets of relationships both to the original book and to other digital representations of it.

[11] By now, anyone familiar with the scholarship of Jerome McGann, and particularly with his famous essay "The Rationale of Hypertext", will have noted the indebtedness of my argument to his. In particular, my discussion of the "one-to-one" relationship of description to original is an intentional echo of McGann's argument that "the facsimile edition stands in a one-to-one relation to its original." I would like to turn now to an even more direct engagement with "The Rationale of Hypertext" through an analysis of how the relationship of the codicological description to the artifact it describes has changed in that formerly both tended to be in codex form, and thus to utilize similar technologies—e.g. indexes, glossaries, and concordances—whereas in a digital environment a description lacks such symmetry of form with the object it describes. In his essay, McGann articulates the many difficulties frequently encountered in using printed critical editions to study other printed books:

"Brilliantly conceived, these works are nonetheless infamously difficult to read and use. Their problems arise because they deploy a book form to study another book

form. This symmetry between the tool and its subject forces the scholar to invent analytic mechanisms that must be displayed and engaged at the primary reading level— e.g. Apparatus structures, descriptive bibliographies, calculi of variants, shorthand reference forms, and so forth."

Printed codicological descriptions are subject to the same limitations—and the same "crucial problem"—as critical editions, and those who have labored to become familiar with the "abbreviated and coded forms" (as McGann terms similar features in printed critical editions) and collational formulae of such descriptions will attest to the difficulty of their use.

[12] As a footnote to this discussion, it is worth noting that the existence of collational formulae as we use them points to the capacity of the printed book to shape—and perhaps misconstrue— our conception of the manuscript book, not merely in deploying "a book form to study another book form", but in invisibly shaping what our notion of a book is in the first place. Consider the following from Fredson Bowers's *Principles of Bibliographical Description*:

The collational formula and the basic description of an edition should be that of an ideally perfect copy of the

original issue. A description is constructed for an ideally perfect copy, not for any individual copy, because an important purpose of the description is to set up a standard of reference whereby imperfections may be detected and properly analyzed when a copy of a book is checked against the bibliographical description. In a very rare book the evidence may not be sufficient to construct a perfect description, but it is better to aim at this perfect description, even though its collational formula may be incomplete and full of queries, than to misrepresent a book by describing only an imperfect individual copy. (113)

But of course in the world of manuscript books, there is only the "very rare book", the "imperfect individual copy". No "standard of reference" is possible in a set of one, nor can we speak of "an edition" of a manuscript book. This should serve as a caution, then, against applying the principles and practices of describing printed books too liberally to those of describing manuscript books.

[13-14]
The rubrication, historiated initials, and foliated borders of incunables remind us that in the early days of print the concept of what a book should be was dominated by the manuscript codex. During recent centuries, the opposite is true; descriptions of manuscript books bear

witness to the dominance of printing in forming our collective notion of what a book should be, and thus we have, for example, assigned them titles and expressed their structures in collational formulae that better reflect the realities of printed rather than manuscript books. As we seek to liberate our codicological descriptions from the constraints of "being compelled to operate in a bookish format," we should also bear in mind the opportunity to correct the assumption that such books operate—and should be described—in parallel with printed books. Both our tools and our mindsets need to be liberated from print if we are to achieve accurate representations of artifacts that were produced before the advent of printing.

Our ideal for original artifacts—the manuscript codices themselves—is that they remain as stable and fixed in time as possible, the goals of our best curation and conservation efforts. But we should be eager to escape the fixity of our tools for working with and describing manuscript books—tools that are often byproducts of the technologies of the printed codex—and embrace instead new purposes and uses for our codicological descriptions, complex new sets of relationships between books, their surrogates, and the technologies we develop to study both, and our opportunity to move beyond the book in order to understand it better.