Adding Value to the University of Oklahoma Libraries History of Science Collection through Digital Enhancement

Maura Valentino

"In getting my books, I have been always solicitous of an ample margin; this not so much through any love of the thing itself, however agreeable, as for the facility it affords me of penciling suggested thoughts, agreements and differences of opinion, or brief critical comments in general."

—Edgar Allan Poe

ABSTRACT

Much of the focus of digital collections has been and continues to be on rare and unique materials, including monographs. A monograph may be made even rarer and more valuable by virtue of hand written marginalia. Using technology to enhance scans of unique books and make previously unreadable marginalia readable increases the value of a digital object to researchers. This article describes a case study of enhancing the marginalia in a rare book by Copernicus.

BACKGROUND

The University of Oklahoma Libraries History of Science Collections holds many rare books and other objects pertaining to the history of science. One of the rarest holdings is a copy of Nicolai Copernici Torinensis *De revolvtionibvs orbium coelestium* (*On the Revolutions of the Heavenly Spheres*), *libri VI*, a book famous for Copernicus' revolutionary astronomical theory that rejected the Ptolemaic earth-centered universe and promoted a heliocentric, sun-centered model.

The History of Science Collections' copy of this manuscript contains notes added to the margins. Similar notes were made in eight different existing copies, and the astrophysicist Owen Gingerich determined that these notes were created by a group of astronomers in Paris known as the Offusius group.[1] The notes are of significant historical importance as they offer information on the initial reception of Copernicus' theories by the Catholic community. Having been created almost five hundred years ago in 1543, the handwriting is faded and the ink has absorbed into the paper.

Maura Valentino (maura.valentino@oregonstate.edu) is Metadata Librarian, Oregon State University, Corvalis, Oregon. Previously she was Digital Initiatives Coordinator at the University of Oklahoma.

Written in cursive script, the letters have merged as the ink has dispersed, adding to the difficulties inherent in reading these valuable annotations.

The book had previously been digitized, and while some of the margin notes were readable, many of the notes were barely visible. Therefore much of the value of the book was being lost in digital form. To rectify this situation the decision was made to enhance the marginalia. It was further decided that once the margin notes were enhanced, two digital representations of each page that contained notes would be included in the digital collection. One copy would present the main text in the most legible fashion (figure 1) and the second copy would highlight the marginalia and ensure that these margin notes were as legible as possible even if in doing so the readability of the main text was diminished (figure 2).
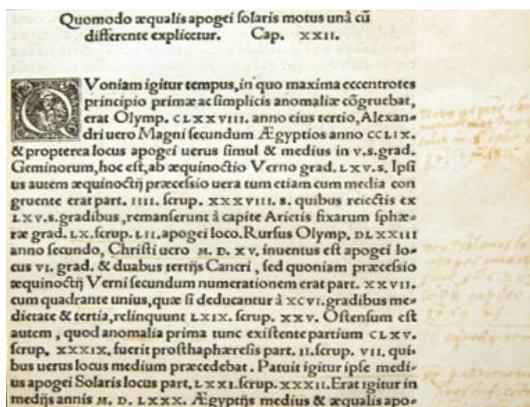
Figure 1. Text readable.
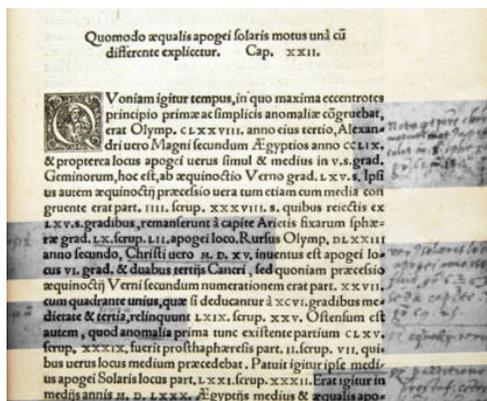
Figure 2. Marginalia enhanced.

While creating a written transcript of the marginalia was considered and would have added some value to the digital object, this solution was rejected in favor of digital enhancement for the following reasons. Many of the notes contained corrections with lines drawn to the area of text that was being changed, or bracket numbers (figure 3). In addition, some of the notes are corrections of numbers or tables, so a transcript of the text would do little to demonstrate the writer's intentions in creating the margin note (figure 4).

Figure 3 .Bracketed corrections.

Figure 4. Numerical corrections.

Also, sometimes there was bleed through from the reverse page, further disrupting the clarity of the marginalia (figures 5 and 6). Therefore it was determined that making the notes more readable through digital enhancement would provide the collection's users with the most useful resource.
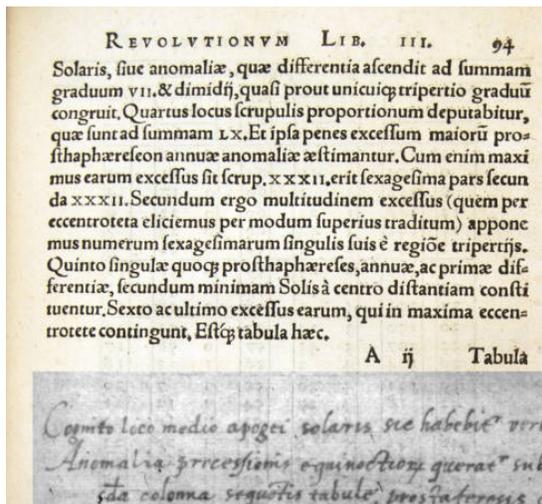
Figure 5. Highlighted—bleed through reduced
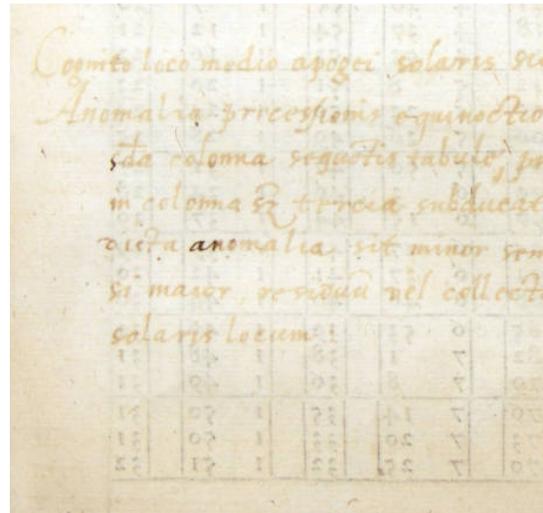
Figure 6. Bleed through behind. marginalia.

The book can be viewed in its entirety here:
http://digital.libraries.ou.edu/cdm/landingpage/collection/copernicus

LITERATURE REVIEW

"Modification of photographs to enhance or change their meaning is nothing new. However, the introduction of techniques for digitizing photographic images and the subsequent development of powerful image editing software has both broadened the possibilities for altering photographs and brought the means of doing so with the reach of anyone with imagination and patience."[2]

—Richard S. Croft

The primary goal of this project was to give researchers in the history of science the ability to clearly decipher the marginalia created by the astronomers of the Offusius group as they annotated the book using the margins as an editing space. The literature agrees that marginalia is an important piece of history worth preserving. Hauptman states, "The thought that produces the necessity for a citation or remark leads directly into the marginal notation."[3] He also adds, "Their close proximity to the text allows for immediate visual connection."[4] Howard asserts, "For writers and scholars, the margins and endpapers became workshops in which to hammer out their own ideas, and offered spaces in which to file and store information."[5] She also adds that marginalia can "serve as a form of opposition."[6] This is true in this case as some of the marginalia

contradicts Copernicus. Nikolova-Houston argues for the historical aspect: "Each of the marginalia and colophons is a unique production by its author, and exists in only one copy."[7] She goes on to add, "Manuscript marginalia and colophons possess historical value as primary historical sources. They are treated as historical evidence along with other written and oral traditions."[8] Such ideas provide a strong justification for the implementation of marginalia enhancement in digital collections.

As mentioned above, it was determined that a transcription would not have had the same effect as digital enhancement of the margin notes. This approach is also supported by the literature. For example, Ferrari argues for the digital publication of the marginalia that Fernando Pessoa, the Portuguese writer, made while reading. One of the cornerstones of his argument is that digital representation of marginalia allows the reader not only to see the words but also the underlining and other symbols that are not easily put into a transcript. In this way, the user of the digital collection obtains a more complete view of the author of the marginalia's intent.[9]

Another goal of this project was the general promotion of the University of Oklahoma's History of Science Collections. Johnson, in his *New York Times* article, notes that marginalia lend books an historical context while enabling users to infer other meanings from their texts.[10] He also quotes David Spadafora, president of the Newberry Library in Chicago, who proclaims that "the digital revolution is a good thing for the physical object." As more people access historical artifacts in electronic form, he notes, "The more they're going to want to encounter the real object."[11] In this way, enhancement of the marginalia in digital collections can lead to further exposure for the collection and to greater use of the physical objects themselves.

Using digital enhancement is not a new idea. Morgan asserts, "The innovation of the World Wide Web is its exciting capacity for space that, while not limitless, is weightless and far less limited that that of the printed book."[12] Le, Anderson and Agarwala also add, "Local manipulation of color and tone is one of the most common operations in the digital imaging workflow."[13]

The literature shows that other projects have used enhancement of the digital object to increase the usefulness of the original artifact. One of the projects pursued during the Library of Congress's American Memory initiative involved the digitization of the work of photographer Solomon Butcher. In this case, technicians were able to enhance an area of one photograph that was blurry in normal photographic processes and allow the viewer to see inside a building.[14] The Archivo General de Indias also used digital enhancement to remove stains and bleed-through from ancient manuscripts and render previously unreadable manuscripts readable.[15] In an article advocating for a digital version of William Blakes's poem *The Four Zoas*, Morgan notes that some features of the manuscript can only be seen in the digital version rather than a transcription: "Sections of the manuscript show intense revision, with passages rubbed out, moved earlier or later in the manuscript, and often, added in the margins."[16]

Digital processing is not limited to the use of photo editing software. Although Giralt asserts that it is a common method, "the ample potential for image control and manipulation provided by digital technology has stirred a great interest in postproduction, and digital editing."[17] Other projects have used various technologies to enhanced images to give added meaning to a digital image. Once again, in her article advocating for the digitizing of William Blake's *The Four Zoas*, Morgan asserts that various enhancement technologies would help readers obtain the greatest benefit from the manuscript. For example, providing "the added benefit of infra-red photography," would allow "readers to see many of the erased illustrations."[18] She even hopes coding will enhance the usefulness of a digital object: "Our impulse to use XML in order to richly encode a text works against passivity. With coding we clarify a work down to its smallest units, and illuminate specific aspects of its structure, aspects that are often less obvious when the work is presented in the form of a printed book."[19]

METHOD

Locating the Marginalia

Each page of the book had been previously scanned and was stored in Tagged Image File Format (TIFF). Each digital page (TIFF image) was carefully examined for marginalia. This was achieved by examining the image in Adobe Photoshop using the Zoom Tool to enlarge the image as necessary. As many notes were barely visible, the entirety of each page had to be examined in detail to ensure that margin notes were not overlooked. Enlargement of the image in Photoshop greatly facilitated this process.

Enhancing the Marginalia

Once all the pages with marginalia were identified, each page was loaded into Adobe Photoshop for digital processing and enhancement. The following procedure was used (*Note: The specific directions that follow reference Adobe Photoshop CS4 for Windows but can be generally applied to most software programs intended for photo editing*):

1. Using the Zoom Tool, the image was enlarged to facilitate examination and interaction with the marginalia.
2. Individual margin notes were selected using the Rectangular Marquee Tool. The area selected included any lines that were drawn from the notes to the original text so it would be clear to what text the margin note referred.
3. As the handwritten margin notes were orange in tone, a blue filter was applied (as blue is the contrasting color to orange) by selecting Adjustments from the Image menu and then choosing Black and White to display the Black and White Dialog box. In the Black and White dialog box, Blue Filter was selected from the Preset drop-down menu. This small adjustment greatly enhanced the readability of the margin notes.
4. With the area still selected, Adjustments was again selected from the Image menu. From that Adjustments submenu, Brightness and Contrast was selected. Adjustments were made to both these values using the sliders presented by the resulting dialog box to

further enhance the margin notes legibility. For this particular project, the values selected were generally negative twenty for contrast and positive twenty for brightness.

File Naming Conventions

Each enhanced image was saved with the same filename as the digital image of the original manuscript page, but with an A (for annotated) added to the end of the filename. This naming scheme enabled a distinction between pages with and without enhanced marginalia. This series of steps was repeated for each page (see table 1).

Page Name	Explanation
Book Spine	Pictures of the covers
Book Cover	
Inside Cover	
Blank Page	With Ruler to Measure Page
Folio - 001	Page 1 as originally scanned
Folio - 001 Verso	Page 1, reverse side, as originally scanned
Folio - 001 Verso A	Page 1, reverse side, with highlighted marginalia
Folio - 002	Page 2 as originally scanned
Folio - 002A	Page 2 with highlighted marginalia
Folio - 002 V	Page 2, reverse side
Folio - 002 V A	Page 2, reverse side, with highlighted marginalia

Table 1. Filenames.

Importing into the Digital Management System

CONTENTdm was the digital management system selected for this project. All original manuscript page images and enhanced marginalia page images were imported into CONTENTdm following their creation. The next step was to bring all the pages into CONTENTdm as one compound object. A Microsoft Excel spreadsheet was created with a line for each page, annotated or not. Only three fields were used: title, rights, and filename. A description of the book was placed on its History of Science Digital Collections webpage with a link to the compound object in CONTENTdm, so further metadata was not necessary and can always be added later. The first row only contained the title of the book (no filename). There were tiffs available of the cover, the bookend, the inside cover, and the book with a ruler. These were the next rows. Then we began with the pages and titled them as the pages were numbered. There were ten pages numbered with Roman numerals and then the pages began with alphanumeric page numbers. Each page that had handwritten notes had the original page (page 2, for example) and the page with the

notes highlighted (page 2 Annotated). This would allow the viewer to view the pages in their original form or with the notes highlighted or both, depending on each user's research interests. Once the Excel file was complete with each page and its filename entered in a row, the file was saved as a tab-delimited file.

Import into CONTENTdm required that all the TIFF files be in one folder. Once the files were moved, the CONTENTdm Compound Object Wizard was used to import. This book was imported as a compound object with no hierarchy. As this book was published in 1593, it has no chapters. To specify page names, the choice to label pages using tab-delimited object for printing was used. The filenames did not contain page numbers, and the choice to label pages in a sequence was not an option, as two copies of each annotated page existed.

Each object imported into CONTENTdm has a thumbnail image associated with it. CONTENTdm will create this image, but the cover of this book is not attractive, so a JPEG file was created using an image from the book that is often associated with Copernicus (see figure 3).

CONCLUSIONS

This project resulted in a digital representation of the physical book that is much more useful to researchers than the original, unenhanced digital object. This History of Science Collection holds not only the first edition of books important to the history of science, but the subsequent editions so that researchers can see how the ideas of science have changed over time. This new digital edition of *De Revolutionibus* allows researchers to see how another scientist made corrections in Copernicus' book as one step in the change in theory over time and insight into the reaction of the Catholic Church.

The format that CONTENTdm creates for the object and a clear naming scheme allow the user to view the pages with or without the marginalia, thus making this a useful object for many types of users (see figure 4).

However, using Photoshop to highlight areas of a page allowed the digital initiatives department to understand the power of this tool.

In understanding the utility and power of Photoshop, the digital initiatives department has determined it to be a useful tool in other projects. A project to eliminate some images of people's fingers that inadvertently were photographed along with pages in a book or manuscript has been added to the queue. In future, digitized books or manuscripts with useful notes will undergo these enhancement processes.

REFERENCES

1. Owen Gingerich, "The Master of the 1550 Radices: Jofrancus Offusius," *Journal for the History of Astronomy* 11 (1993): 235–53, http://adsabs.harvard.edu/full/1993JHA....24..235G.

2. Richard S. Croft, "*Fun and Games with Photoshop: using Image Editors to change Photographic Meaning*" (In: Visual Literacy in the Digital Age: Selected Readings from the Annual Conference of the International Visual Literacy Association (Rochester, NY October 13-17, 1993)): 3-10.

3. Robert Hauptman, *Documentation: A History and Critique of Attribution, Commentary, Glosses, Marginalia, Notes, Bibliographies, Works-Cited Lists, and Citation Indexing and Analysis* (Jefferson, NC: McFarland, 2008).

4. Ibid.

5. Jennifer Howard, "Scholarship on the Edge," *Chronicle of Higher Education* 52, no. 9 (2005).

6. Ibid.

7. Tatiana Nikolova-Houston,"Marginalia and Colophons in Bulgarian Manuscripts and Early Printed Books," *Journal of Religious & Theological Information* 8, no. 1/2, (2009), http://www.tandfonline.com/doi/abs/10.1080/10477840903459586#preview.

8. Ibid.

9. Patricio Ferrari, "Fernando Pessoa as a Writing-Reader: Some Justifications for a Complete Digital Edition of his Marginalia," *Portuguese Studies* 24, no. 2 (2008): 69–114, http://www.jstor.org/stable/41105307.

10. Dirk Johnson, "Book Lovers Fear Dim future for Notes in the Margins," *New York Times*, February 20, 2011, http://www.nytimes.com/2011/02/21/books/21margin.html?_r=3&emc=tnt&tntemail1=y&

11. Ibid

12. Paige Morgan, "The Minute Particular in the Immensity of the Internet: What Coleridge, Hartley and Blake can teach us about Digital Editing," *Romanticism* 15, no. 2 (2009), http://www.euppublishing.com/doi/abs/10.3366/E1354991X09000774.

13. Y. Li, E. Adelson, and A. Agarwala, "ScribbleBoost: Adding Classification to Edge-Aware Interpolation of Local Image and Video Adjustments," *Eurographics Symposium on Rendering* 27, no. 4 (2008), http://www.mit.edu/~yzli/eg08.pdf.

14. S. Michael Malinconico, "Digital Preservation Technologies and Hybrid Libraries," *Information Services & Use* 22, no. 4 (2002): 159–74, http://iospress.metapress.com/content/gep1rx9rednylm2n.

15. Ibid.

16. Morgan, "Minute Particular."

17. Gabriel F. Giralt, "Realism and Realistic Representation in the Digital Age," *Journal of Film & Video* 62, no. 3 (2010): 3,
    http://muse.jhu.edu/journals/journal_of_film_and_video/v062/62.3.giralt.html.

18. Morgan, "Minute Particular."

19. Morgan, "Minute Particular."