

A NEW MODEL FOR IMAGE-BASED HUMANITIES COMPUTING

A Thesis

by

JACOB HOHMANN BROWN

Submitted to the Office of Graduate Studies of  
Texas A&M University  
in partial fulfillment of the requirements for the degree of

MASTER OF ARTS

August 2008

Major Subject: English

A NEW MODEL FOR IMAGE-BASED HUMANITIES COMPUTING

A Thesis

by

JACOB HOHMANN BROWN

Submitted to the Office of Graduate Studies of  
Texas A&M University  
in partial fulfillment of the requirements for the degree of

MASTER OF ARTS

Approved by:

Chair of Committee,	Maura Ives
Committee Members,	Craig Kallendorf
	Eduardo Urbina
Head of Department,	M. Jimmie Killingsworth

August 2008

Major Subject: English

## ABSTRACT

A New Model for Image-based Humanities Computing. (August 2008)

Jacob Hohmann Brown, B.A., University of Dallas

Chair of Advisory Committee: Dr. Maura Ives

Image-based humanities computing, the computer-assisted study of digitally-represented “objects or artifacts of cultural heritage,” is an increasingly popular yet “established practice” located at the most recent intersections of humanities scholarship and “digital imaging technologies,” as Matthew Kirschenbaum has pointed out. Many exciting things have been and are being done in this field, as multifaceted multimedia projects and “advanced visual and visualization tools” continue to be produced and used; but it also seems to lack definition and seems unnecessarily limited in its critical approach to digital images. That is, the textual mediation required to make images usable or knowable, and the kinds of knowledge images offer, often goes unexamined, and the value of creative or deformative responses to images overlooked. This thesis will suggest Blake’s production of the *Laocoön* as a model for a more open and relevant approach to images, will analyze what image-based humanities computing does and how Blake’s engraving recapitulates these actions, and will describe how acritical approaches to image description could be integrated and used, and how images could function as graphic mediation for other materials, in this field.

Blake’s idiosyncratic *Laocoön* exemplifies the ways that creators or editors respond to and describe images and the ways they use images to illuminate text. In entitling his plate “[Jah] & his two Sons [...]” and filling it with descriptive text, Blake shares the focus of image-based humanities computing on images as things to be broken down, described, and understood. But Blake’s classification and description, deformative in misreading the image, reveals the true nature of such mediation

and the need for a more open system, one which allows observers to record how they interpret an image, perhaps best accomplished in image-based humanities computing through semantic web technologies like folksonomy tagging or collaborative wiki formats. And Blake's act of pulling a pre-existing image out of context and applying it to a new textual work suggests a new function for images and the highly structured image databases of image-based humanities computing, to clarify or complicate textual works through graphic mediation.

## ACKNOWLEDGMENTS

I want to thank especially my committee chair and research advisor, Dr. Maura Ives, without whom I could not have written this thesis, and without whose course on digital editing I would not have wanted to. Her experience in, and knowledge of, the digital humanities, as well as her constant feedback on my work, were invaluable in writing this thesis.

I would also like to thank my committee members, Dr. Craig Kallendorf, whose course on the histories of the book fueled my interest in bibliography and illustration studies, and Dr. Eduardo Urbina, whose work in the *Cervantes Iconography Project* is inspiring, for agreeing to serve on my committee.

My gratitude also goes to all the faculty and staff of the English Department at Texas A&M University, all of whom have been immensely helpful, and especially to the Director of Graduate Studies, Dr. Sally Robinson, and to Ms. Paulette Leshner.

Finally, I would like to thank my fiancée, Elizabeth Deacon, without whose love and support I would surely have failed, and whose naïve yet perceptive response to the system for “self-illustrating” texts described in Section IV—“So, you’re trying to put [illustrators] like me out of business?”—just made me laugh.

## TABLE OF CONTENTS

CHAPTER	Page
I INTRODUCTION . . . . .	1
II IMAGE-BASED HUMANITIES COMPUTING . . . . .	14
III NEW FORMS OF IMAGE EKPHRASIS . . . . .	31
IV NEW FORMS OF TEXT ILLUMINATION . . . . .	41
V SUMMARY . . . . .	52
WORKS CITED . . . . .	56
VITA . . . . .	62

## LIST OF FIGURES

FIGURE	Page
1	William Blake’s <i>Laocoön</i> , Copy B, from <i>The Blake Archive</i> . . . . . 2
2	The University of Victoria Image Markup Tool, used here to remix an image and editorial annotation from the <i>Cervantes Iconography Project</i> (Pinelli). . . . . 20
3	“[Inote] functions most powerfully when used in conjunction with the Archives image searching capabilities, where it can open an image found by the search engine, zoomed to the quadrant of the image containing the object(s) of the search query, with the relevant textual annotation displayed in a separate window.” (“Technical Summary”). . . . . 23
4	The Iconclass hierarchy for the action of the <i>Laocoön</i> , as represented in the <i>Iconclass Libertas Browser</i> . . . . . 25
5	Screenshot of <i>Woman Aircraft Worker [...]</i> , with annotations, from Flickr.com. . . . . 37
6	“they carried him up onto the roof [...] There were chimney swifts in the sky.” . . . . . 48
7	“they went into the Duomo and prayed. It was dim and quiet [...].” 49
8	“When they had to say goodbye, in the station at Milan [...].” . . . 50
9	“while riding in a taxicab through Lincoln Park [...].” . . . . . 51

## CHAPTER I

## INTRODUCTION

William Blake's late plate commonly referred to as the *Laocoön* is an odd creation, even for the idiosyncratic poet-artist. As one of his works of "illuminated printing," the *Laocoön* mingles text and image in typically Blakean ways, but this plate is different: the text is apoetic and often disconnected, and much of it is wrapped around, or crammed into the empty spaces of, the illustration, while the illustration itself is an unnaturally careful depiction of a historical art object, the Hellenistic sculpture representing the Trojan priest strangled by sea-serpents along with his two sons. Blake himself, however, did not call this work *Laocoön*: he seems rather to have entitled it "[Jah] & his two Sons Satan & Adam as they were copied from the Cherubim of Solomons Temple by three Rhodians & applied to Natural Fact or History of Ilium," certainly an original interpretation of the piece.<sup>1</sup> The text of the plate perhaps elucidates this historical re-reading and covers a variety of other topics including religion, war, money, and art, and is, as Irene Taylor notes, "clearly a kind of summary index to Blake's later thought" (qtd. in James: 228). These inscriptions are also clearly visually oriented: the text is often positioned in such a way that it is obviously intended to identify or relate to parts of the illustration—thus one serpent is represented as "Evil" and the other as "Good"—and Blake uses different scripts and non-English characters, and experiments with different kinds of lines as punctuation,

---

This thesis follows the MLA citation style.

<sup>1</sup>This title is located prominently below the sculpture's plinth. It is hard to tell, but Blake seems earnest in his belief that the sculpture was taken out of its true context when used to represent the Trojan priest. Thus, while Blake appears to reject allegory in restoring the true meaning of the image (James 229), he is—probably—doing precisely what he claims the "three Rhodians" had done.





is, understanding this image (in the most basic sense) involves observation and visual study, but also textual and even critical mediation. Thus, for example, *The Blake Archive*, a hypermedia collection of Blake's works, offers a digital facsimile of the plate, but also attempts a comprehensive textual description of the work: the central figure of Laocoön or God, the editors point out, "is bearded and his hair is long. His mouth is slightly open and his visage expresses great pain. His genitals are clearly pictured. He faces forward, with his right leg bent at the knee and his left leg extended to the right," and so on ("Illustration Description").<sup>2</sup> Transcription for the plate is a little less straightforward, due to the "spatial and graphic qualities" of the text, but the editors of *The Blake Archive* manage it well, beginning with "If Morality was Christianity Socrates was the Saviour" at the base of the statue, and proceeding in a generally widening and clockwise gyre to "All is not Sin that Satan calls so / all the Loves & Graces of Eternity / Drawn & Engraved by William Blake," located on the plinth ("Textual Transcription"). As far as interpretation goes, David James convincingly argues, in his analysis of the plate, "Blake's *Laocoön*: A Degree Zero of Literary Production," that the *Laocoön* is an "attack on the role of money in reifying art and turning it into a commodity," citing the failure of Blake's earlier "commercial engraving projects," the plate's unusual and condemning emphasis on money—for example, "Christianity is Art / & not Money / Money is its Curse"—and the fact that only a few copies of the plate were printed, thus denying the plate "existence in the economic and political transactions of a fallen world" (231, 234).

But perhaps the most interesting aspect of the work is the strange relationship between text and image—especially in how the title and inscriptions relate to and complicate the engraving of the *Laocoön*, and vice versa. While the meaning of the

---

<sup>2</sup>And it does go on: a picture here is apparently worth 1227 words, the total word count for the "Illustration Description."

relationship here between text and image is not quite obvious, this relationship is obviously important. As James points out, the arrangement of text and image “precludes a sequential linear reading of the inscriptions” (228), and reading here proceeds somewhat haphazardly as the eye is drawn from one part of the figures to another—the plate requires almost constant transition between observation (or image analysis) and reading.<sup>3</sup> Blake made his engraving of the sculpture sometime around 1815 as a commissioned work for an encyclopedia article on sculpture, but the inscriptions were “almost certainly” added much later, in 1826 or 1827, when Blake made a few copies—it was probably never offered for sale—of the plate (James 226; Eaves et al., “Copy Information”). Thus when Blake returned to the engraving in the mid 1820s, he formed a new interpretation for what it depicted: this engraving reflected no Il-ionic event—or only secondarily so—it showed, rather, Jehovah, the “serpent-holder,” along with Satan and Adam, entangled with a couple of serpents. This new reading Blake fleshed out with inscribed references to “Good & Evil,” “Hebrew Art,” Lilith (in Hebrew), and so on, which are tied to the image, while even the more removed inscriptions—like “Spiritual War” or “He repented that he had made Adam / of the Female, the Adamah”—seem entailed by the image’s perceived content. Although, as James dryly notes, nowhere in Blake’s other work “do we find Jehovah, Satan, and Adam struggling with snakes” (229), Blake clearly derives some mythological or theological meaning from this image, meaning which radiates outward from the image, and which is mediated and explicated by the textual inscriptions across the plate. While the title and inscriptions relate to the image in highly idiosyncratic and original ways, they offer one—admittedly deformative—reading of the image, and they reflect one of the image’s “possibilities of meaning.”

---

<sup>3</sup>Or it requires that the text be read as an image is “read,” “all at once,” as David Erdman points out (qtd. in James: 228).

At the same time, though there is no other instance in Blake's work of Jehovah battling serpents, as James observes, the "cognate incident" represented by the sculpture does exist elsewhere—or everywhere—in his work, and James argues that this engraving represents "the reversal of the fall, as specifically figured in the setting of the 'Limit of Opacity' and the 'Limit of Contraction' [described] most fully in *The Four Zoas*," and that "imagination-God-Jesus," identified here as Jehovah, "is projected into the forms (Satan and Adam) in which it must enter space and time and confront materiality and reason" (229–230). That is, rather than having effected a new theological truth, this engraving seems to have been chosen by Blake to illuminate an incident—or at least an idea—from his "labyrinthine and frequently self-contradictory mythic theology" (229). And the inscriptions scrawled across the plate, though they are certainly related to the content of the image, are "a kind of summary index" to Blake's thought, as has been noted, and no doubt existed in some form before the plate was inscribed; and they seem to make use of the engraving more as a kind of visual guide to this index. Blake then seems to have had this incident and these thoughts in his mind—or scattered around his workshop—when he came across the *Laocoön* engraving again in 1826 or 1827, and this engraving must have seemed an ideal (or at least appropriate) illustration for representing this incident, and an ideal (or at least adequate—intaglio plates were not cheap) place for relating these inscriptions. Though the appropriateness of the engraving as illustration is questionable, it does offer one graphic expression of Blake's ideas about the "reversal of the fall," and it does reflect one of this text's "possibilities of meaning."

Of course, these hypotheses—that the image of the *Laocoön* came first and is supplemented by textual mediation and that the textual idea came first and is supplemented by visual mediation—are not mutually exclusive: inspiration and creativity are not so neat and clean, and the final plate is probably the result of some kind of

“feedback loop” between text and image. But such analysis reveals the ways that texts and images (are made to) interact. Thus, in a kind of creative ekphrasis, text may be derived or transmediated from an image to represent or perform that image, though in attaching itself to the image, this text may also complicate or deform it: that is, in entitling his plate “[Jah] & his two Sons Satan & Adam [. . .],” Blake adds meaning to the image that the image may or may not sustain. But without his title and inscriptions, the engraving represents simply the Trojan priest and his two sons—or, even more simply, three guys with a couple of snakes. On the other hand, in a kind of critical or interpretive ekphrasis, textual description, transcription, and interpretation may also be derived from and intended to represent an image, and by attaching itself to the image may clarify or aid access to the image—though it may also deform the image, a possibility played down by archives and scholars. At the same time, an image, whether picked out from pre-existing work and matched with a text, as Blake chose the *Laocoön* to represent his ideas about the “reversal of the fall,” or developed in response to a text, as illustration has typically been done, may illuminate that text, by providing some visualization of its content, but may also complicate it: that is, whatever Blake’s ideas were about “the reversal of the fall,” or at least about Jehovah, Satan, and Adam, before the inscription of the plate, they were ever after colored by this serpentine struggle.

Images like Blake’s *Laocoön* are things to be not only represented—that is, made accessible with relatively little mediation—but also described, mediated, and interpreted; and image-based humanities computing, the computer-assisted study of digitally-represented graphic or multimedia objects, allows new ways—or perhaps improves old ways—of interacting with images, and has typically focused on critical ekphrasis: on extracting information or data from such images, on organizing them as “structured data,” and on analyzing what an image—or image collection—means.

And such is the beginning of wisdom—or at least of some form of image-based knowledge. While observation may be the ideal way to understand an image (in the most basic sense), the image rarely reveals its every aspect through simple observation. As A. A. Goodrum, B. C. O’Connor, and J. M. Turner note in their introduction to a special issue of *Computers and the Humanities* devoted to digital images (1999), “the same structure and nature of these [visual] documents that has rendered them congenial [to a culture] has made comparison, synthesis, and analysis very difficult” (291), and while images invite description, mediation, and interpretation, the way these interpretive actions proceed is not straightforward nor is their relevance and authority guaranteed. However, as Julia Flanders suggests in “Trusting the Electronic Edition,” it is only through such description and analysis that the meaning and importance of an image is revealed: though the image is “so detailed that one could never describe it with sufficient accuracy,” it is through such delimiting description and analysis that we discover what graphically-derived information “we really can and do use” (308). That is, images are at once opaque and full of meaning, and textual or editorial mediation both reverses this opacity and limits the image’s meaning to something knowable. And image-based humanities computing is clearly yet paradoxically textually oriented: not only is such description and analysis—as well as the computer markup underlying the field—generally performed in and through text, but bringing images “into the fold of structured data” means making them more like “searchable [or machine-readable] text” (Flanders 303, 308). Even more, this field is actually founded on textual description and analysis: because digital images are opaque and are not searchable like digital text, they require a kind of “pre-processing,” in which textual or editorial mediation is derived from and attached to (or associated with) an image, so that the image can be critically known and correctly used.

However, in their concern for accuracy, authority, and authenticity, many scholars

engaged in image-based humanities computing fail to realize the inherently deformative nature of critical representation, description, and analysis, as well as the critical possibilities of embracing and integrating this deformance. As Jerome McGann notes in “Visible and Invisible Books: Hermetic Images in N-Dimensional Space,” all critical representation “deforms its object,” because “the object perpetually shifts and mutates under the influence of its perceivers,” and involves only “a certain perspective on the object,” one which is perceptive insofar as it understands its “own powers and limitations” (*Radiant Textuality* 287). Even more, description and analysis provide only a certain perspective, not only because the image contains too much information to be fully described (cf. Flanders 308), but also because different people will see different things in an image—thus the central figure of the *Laocoön* may be a Trojan priest, Jehovah, a man, or even a character in one’s head. And there are certainly cases in image-based humanities computing in which the perspective of the creator or editor does not sufficiently reflect the image’s form and content, or in which their perspective does not align with the user’s perspective. But though poor description is often due to people’s lies, laziness, stupidity, or lack of self-consciousness (Doctorow), such misalignments (especially if intelligently designed) may be valuable. While Blake’s identification of the *Laocoön* as depicting Jehovah, Satan, and Adam, may be considered inaccurate, it also reflects one revealing perspective of the image. If Blake had been an “electronic editor” and tagged the engraving as depicting these figures, then users looking for the Trojan priest may not have discovered this image, it is true, and users looking for Jehovah may have been immensely confused by this image. But this confusion would be mind-expanding: *How is this image a depiction of Jehovah, Satan, and Adam?* Attempting to answer this question is immensely productive, and involves making connections in non-traditional, atypical ways, by analyzing semantic aspects of the image that could work in more ways (or have more meanings) than

one. Images are multifaceted and, to use Edward Tufte's term, "multivariate" (qtd. in McGann: 289), and can sustain—and may need, to reveal all of their potential aspects—a variety of descriptions or identifications, including those which may be deformative or typically (traditionally) exclusive.

But images also function in image-based humanities computing as the mediation, either for their own represented objects or for other—often textual—materials. As Joan Schwartz, following James Elkins and Richard Lanham, notes in "Negotiating the Visual Turn: New Perspectives on Images and Archives," our culture has become increasingly visual, both the quantity and the kind of our culture's images are completely different from those of "past centuries," and the role of images in ordinary communication has been readjusted (108); and this shift is reflected by contemporary culture's increasingly primary field of communication, the internet, where even primarily text-centered pages must make use of their visual elements to remain relevant.<sup>4</sup> That is, for some intents and purposes, the digital image is the best way of communicating or conveying not just visual knowledge. Thus Flanders argues that, for textual or literary studies, "documentary evidence" (for example, digital facsimiles of a Hardy manuscript) and even images tangentially related to a text (for example, a photograph of lapis lazuli associated with Yeats' poem) may be useful, may contribute to the production of some kind of knowledge, and may, in fact, be more than the sum of their textual descriptions or counterparts. Documentary materials do so by "multiplying media" (Eaves 212), by substantiating editorial transcription, and by revealing their often meaningful physical features. Digitally represented documents allow "access to rare material" and provide readers with "primary evidence," and by multiplying the media for a textual work, such images may provide "alternate

---

<sup>4</sup>A disturbing aspect of the "Web," according to Flanders (303).



perspectives” on the text (301, 306). In “Computing: What Has American Literary Study To Do with IT?,” Martha Nell Smith expands on the democratizing aspects of such images, as well as the transparency effected by them: not only are documentary images no longer for use “by experts only,” but such widespread observation can correct the inevitable errors in editorial transcription (837, 849).<sup>5</sup> More importantly, however, such documents are revealing in and through their physical features, and Matthew Kirschenbaum points out, in his introduction to a special issue of *Computers and the Humanities* devoted to image-based humanities computing, that literary forms of this field have descended

from the so-called social or materialist theories of textual production advanced by such scholars as D. F. McKenzie and Jerome McGann, [which] privilege the physical primacy of the manuscript, document, or codex, lending semantic weight to paper, cover, binding, typeface, page layout, illustrations, marginalia, and other extra-linguistic features of the “text itself.” (3)

The physical features of all documentary materials are potentially meaningful, and, by their very nature, such documents are “so detailed that one could never describe them with sufficient accuracy” (Flanders 308).

At the same time, even disconnected images may be used to supplement or reveal some aspect of a textual work. Thus Flanders argues that an image of the semi-precious stone in the context of William O’Donnell and Emily Thrush’s hypertext edition of Yeats’ “Lapis Lazuli,” for example, is potentially revealing and “empowering”:

---

<sup>5</sup>See especially her anecdote about a cow with “outstretched hands” and mistranscribed texts (849).

the images which accompany the text act as windows between the real world and the world of the text, as if to say “yes, that really happened—there really was a piece of lapis lazuli.” The fact that there are images of the things that surround the text gives a “see for yourself” air to the contextual materials: the context offered is positioned as helpful additional information, [and the] mode offered to the reader is one of empowerment. (307)

Such a use of digital images centers on “connectedness and copia,” placing the text “at the center of a matrix around which are clustered all kinds of background information,” and by “supplying this wealth of material the editor seems to be trying to furnish the reader’s mind as a habitable place” for the text (306). And Kirschenbaum suggests the value of relevant digital-born images, “visualizations,” of, for example, a user’s “search results” or “the narrative structure of a text,” in image-based humanities computing (“Refining”). But none engaged in this field have suggested doing what Blake seems to do in the *Laocoön*—despite the ready availability of highly structured image databases: pulling relevant pre-existing images out of context based on their identified form and content, and associating them with new materials to illuminate or complicate those (often textual) objects.

In his introduction to image-based humanities computing, Kirschenbaum offers a definition of what image-based humanities computing, this consistently ill-defined yet already “established practice,” does: it brings “advanced visual and visualization tools to bear on the objects and artifacts of cultural heritage,” and it involves building “editions, collections, or self-identified ‘archives’ that place digital images at the center of their interface (and database)” (3). These tools, which range from Inote—a Java-based software tool designed at the Institute for Advanced Technol-

ogy in the Humanities (IATH) to map annotations for an image to a digital image and to allow organizing hyperlinks—to the color histograms of image manipulation programs like Adobe Photoshop—a feature which provides data about the distribution of color in a digital image—offer ways of breaking images down into “structured data” and of extracting data from images, and may be used in the building of these editions or archives or for personal research. And such editions or archives, intended to produce some form of textual or visual knowledge, may center on presenting a multifaceted, multimedia “scholarly edition” of image-based works (as *The Blake Archive* does), or on allowing and improving “access to collections of visual materials” through “image indexing and classification” and an analysis of “image users and image uses,” as Corinne Jörgensen notes in “Access to Pictorial Material: A Review of Current Research and Future Prospects” (293). In a nutshell, image-based humanities computing—as it is currently practiced—involves representing images, attaching metadata to images, breaking them down into discrete semantic parts so that they can be known and used in a variety of ways, linking editorial annotations and images—or images and other images—to facilitate or guide research, identifying all relevant images for all kinds of user queries, or using the computer as computer to practice some form of interpretation—or in some cases deformation—upon images.

This thesis, then, will first describe what image-based humanities computing does, and how it functions, analyzing the editions, archives, projects, tools, and mechanisms of this increasingly popular field of the digital humanities,<sup>6</sup> and identifying where this practice intersects with what Blake, as an image editor, does in the *Laocoön*. It will then offer a new model for the critical ekphrasis or textual me-

---

<sup>6</sup>Dorothy Porter, for example, has pointed out the blossoming interest in “incorporating source images” even in “traditional text encoding projects,” and the proliferation of software tools “focused on image editing and display,” over the last decade (“Examples of Images in Text Editing”)

diation of image-based humanities computing, based on Blake's textual deformance in the *Laocoön*, one which would revalue the idiosyncratic voice and disparate description and make use of technological developments and semantic web features like folksonomy tagging and data-based wikis as more open-ended ways to mediate and understand images. Finally, it will describe a new function for digital images in this field, based on Blake's unusual use of the (pre-existing) *Laocoön* engraving to illustrate his textual work: a system for the automated illumination of text and textual works using the image databases and classification strategies of the field, in which images would be pulled out of context, re-used, and applied to new textual objects to clarify, complicate, or deform them in meaningful ways.

## CHAPTER II

## IMAGE-BASED HUMANITIES COMPUTING

Images, even by themselves, are revealing, especially if beholders know what they are looking at or for—much of their information is conveyed through simple representation and observation; and as McGann suggests, such representation and observation is in itself illuminating: “simply to ease access [to] artistic work enhances our abilities to understand” (“Visible and Invisible Images” 284). Now more than ever, and in our visually-oriented culture, scholars and users are likely to be able to make sense of, to use, and to understand even unidentified images or image collections. And the simple archiving and representation of high-quality images—possible now thanks to Moore’s Law of exponentially decreasing digital memory costs<sup>1</sup>—may facilitate a certain kind of research, and is certainly necessary for any image-based humanities computing project. Even more, it is the computer’s potential for such representation that makes image-based humanities computing such an exciting field; as Kirschenbaum notes, the computer is now a viable “*venue for representation,*” evident in the “genuine excitement of seeing a high-resolution, 24-bit color image wash across the display screen” (“Editor’s Introduction” 4). This excitement, in turn, may lead to productive, visually-oriented questions “about image acquisition and editorial fidelity, not to mention scholarly and pedagogical potential” (4). At the same time, however, the computer, “as computer,” can do far more than simply display images;<sup>2</sup> even more, simple digital representation may be no more than “bravura display,” or function as

---

<sup>1</sup>Flanders identifies the expense of storing digital images (in 1998) as a major point in the argument against image-based humanities computing (301).

<sup>2</sup>Although, as Kirschenbaum has pointed out, it was the computer’s role “as computer,” which lead to skepticism or “apathy towards computers” (4), things seem to have shifted since 2002 (when he wrote his “Editor’s Introduction”), as many scholars have come to realize the value of the computer’s computation.

nothing more than “adornment,” as Flanders points out (309), and without some sort of higher-level information about the image, beholders may be lost, confused, or apathetic.

Metadata—simply “data about data”—offers such higher-level information, and underlies almost every aspect of image-based humanities computing: if images, by themselves, are unstructured data, it is metadata that makes them into (or at least accessible as) structured data. Images by themselves may be revealing, but without metadata, as Howard Besser points out in his *Introduction to Imaging*, the digital image “will be difficult to find or identify and is likely to become useless very quickly.” Besser goes on to define what metadata does, by dividing it into “three broad types,” as follows:

descriptive [metadata], which describes content; administrative, which describes context and form and gives data-management information; and structural, which describes relationships between parts and between digital files or objects. The first is most akin to traditional cataloguing and would describe what an image depicts or its essence. [The second] records information such as how and why a digital object was created, and is used in the management of digital objects. [The third] documents information such as the fact that one image is a detail of the upper-right corner of another image, that a particular image depicts page two of a book of thirty-four pages, or that an image is one item in a series.

Thus image metadata facilitates image searching, by applying basic and even naïve—as well as higher-level—descriptions to the image (thus enabling “end users to find and distinguish between various images”), and provides a scholarly or editorial point of view, through description; it answers questions about “image acquisition and editorial

fidelity” (Kirschenbaum 4) and gives information about file formats, color corrections, compression, revision history, and so on (thus allowing “collection managers to track, preserve, and make accessible digital images”) through its administrative metadata; and it controls the representation of the image (in the edition or archive, for example, and on the screen) through its structural markup. Such metadata also allows images and image collections to be reused and built upon, as Besser notes, and facilitates the necessary migration of image files or image data.<sup>3</sup> Image metadata is clearly integral to image-based humanities computing; thus, in 2001, the Digital Library Forum defined “good” metadata, summarized here by Besser:

[image metadata] is appropriate to the materials digitized and their current and likely use; it supports interoperability; it uses standard controlled vocabularies to populate elements where appropriate; it includes a clear statement on the terms of use of the digital object; it supports the long-term management of digital objects; and it is persistent, authoritative, and verifiable.

That is, such metadata takes into account “image users and image uses,” and would be different, for example, depending on whether it was used in a focused and limited image-based edition or in an archive of varied and disparate images. It facilitates inter-collection operability, doing something like what the Textual Encoding Initiative (TEI) does for text—as Jane Edwards notes, TEI standards “were designed to facilitate data exchange across projects” (343). The use of “controlled vocabularies” assists this interoperability, but also allows for higher-level organization of metadata

---

<sup>3</sup>As Christine Sundt points out in “Digital Projects Past and Present—Survivors or Fossils?,” the survival and continued relevance of image-based computing projects depends almost entirely on digital evolution, on consistently migrating their data to new forms and file formats.

within the project itself. It also clarifies issues regarding use, takes measures against the digital disappearance or obsolescence of its works, and presents the project and its images as authoritative.<sup>4</sup>

This higher-level information is typically attached to the image digitally through some form of textual markup like the Standard Generalized Markup Language (SGML) or its child, the eXtensible Markup Language (XML). While McGann has noted the inherent limitations of using such markup languages, which were designed for “informational or expository” texts and which expose only the “informational content or expository structure” of a work, to describe any kind of imaginative work (“Visible and Invisible Books” 289), and while Kirschenbaum has pondered the perhaps strange absence of image markup using images (as text has been used to mark up text), these markup languages are in fact able to express the data-based aspects (including some forms of description) quite well. While database fields or forms of free text may also be used to attach metadata to images, markup languages like XML support interoperability, higher-level organization, and the ability to link the image file and metadata “inextricably.” As a “hardware- and software-independent format,” as Besser points out, XML allows “information to be structured, stored, and exchanged between what might otherwise be incompatible systems,” and allows users to “generate uniformly structured metadata packages that can be used for sharing,” as well as for migrating data and internet publishing. As a hierarchical and machine-readable language—that is, the computer may be programmed to understand the relationships among

---

<sup>4</sup>As Flanders points out, images are inherently unreliable or deceptively reliable, thanks to image manipulation programs like Adobe Photoshop (303). Hashing or checksum algorithms may be a not especially elegant solution to this problem, but in fact, much of the pressure placed on images to tell the truth has been lifted, as few scholars now view digital facsimiles as complete substitutes for the actual objects; and these projects are rather seen as a place to do useful but primarily preliminary research. This shift may be seen as a loss, but is probably rather a fortunate reflection of the understanding that no facsimile can take the place of the original object.



elements in XML—this format allows complex semantic structuring and searching. It offers the ability for an image to be both “called up” and described in one highly structured document, thus bringing the “parallel universes” of text and image (cf. McGann, *Radiant Textuality* 95) together as one object. XML also works and integrates well with related (and derivative) formats, including the Resource Description Framework (RDF) format—which can describe relationships among XML elements across the Web—and the Scalable Vector Graphics format—which can describe and script graphic structures or elements—both of which are useful in the formation of image metadata.

Image metadata may be attached to an image either globally (to the image as a whole) or locally (to one or more graphic sections of the image). That is, image-based humanities computing may break images down into discrete semantic parts—thus making them more like digital text, where individual ASCII patterns or words can be identified and pulled out—by mapping and identifying exclusive sections of a digital image, and by attaching different kinds of metadata to these sections. These actions involve, in a way, creating a kind of “analytical model” for the layout and substance of the image: though, as Flanders notes, images in general do not yield readily to abstraction (especially of the graphically-oriented parts), and though digital images, which are “inherently resistant” to “electronic inspection, computer-assisted analysis, and retrieval,” reflect and exacerbate this unyielding nature (304), yet such a model is possible and can be immensely useful in understanding an image—for example, as a way of abstracting and analyzing “the formal structure of the text image” (McGann, *Radiant Textuality* 96). The importance of such a model, and especially of image mapping, is revealed in that the first function of one of the first tools of image-based humanities computing, Inote, was to allow “one to build a random set of image points to which one could attach text materials of varying kinds,” as McGann recalls

in *Radiant Textuality* (95). McGann goes on to explain the need for such a function, as well as the traditional and bibliographical inspiration of the editors involved in developing it:

Because information in bit-mapped images cannot be coded for analysis, our technical people were asked if it would be possible to lay an electronic transparency (as it were) over the digital image and then use that overlay as the vehicle for carrying computable marked-up data and hyperlink texts. The idea was to treat the overlay as a kind of see-through page on which one would write text that elucidated or annotated the imaged material “seen through” the overlay. [The idea] originates in scholarly editions that utilize onionskin of other transparent pages to create an editorial palimpsest for complex textual situations. (95)

Though distinctly “hack-like,” this image mapping feature offers creators or editors the precious ability to “point into” and mark up an image file in much the same way that stand-off markup has been used to mark up text.<sup>5</sup> And this function of Inote, which can be observed (albeit primitively) in *The Blake Archive*, where each digital image is divided into quadrants, each of which has its own notes and “illustration description,” has been a common feature of almost all humanities-based image tools, though more recent projects—including the University of Victoria’s Image Markup Tool—have used more sophisticated mechanisms and markup languages, especially SVG, for such image mapping (see Fig. 2 for an example of the Image Markup Tool in action). SVG, an XML file format capable of describing two-dimensional vector shapes, not only allows the mapping and identification of any user-defined image

---

<sup>5</sup>For more information on stand-off markup, see “TEI Standoff Markup Workgroup” on the Textual Encoding Initiative website.

section (the editor could, for example, use SVG to trace an outline around an object within the image and then describe that section in the XML format), but also, as a form of XML, allows this markup to be highly organized and configured within one document, in a “logical structure.”

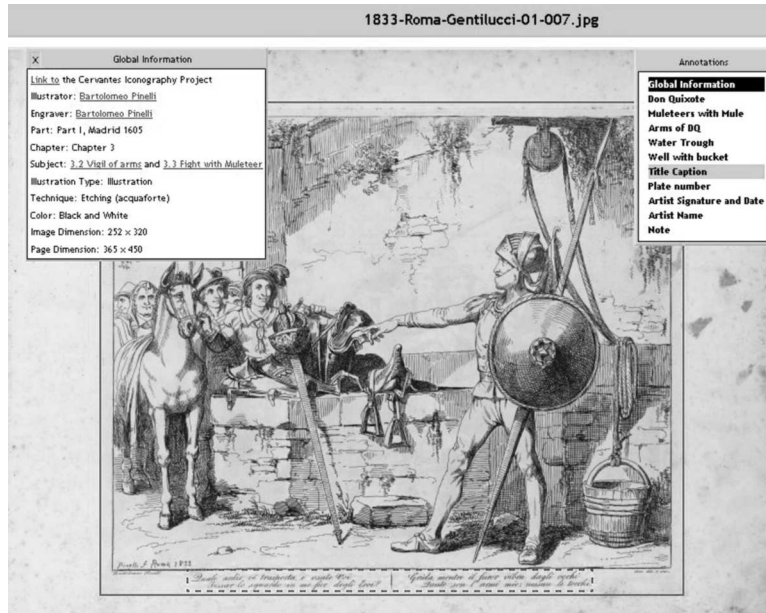


Fig. 2. The University of Victoria Image Markup Tool, used here to remix an image and editorial annotation from the *Cervantes Iconography Project* (Pinelli).

This image mapping and identification enables local (as opposed to global) description and analysis, as well as the description and analysis of the spatial organization of an image. That is, creators or editors may focus on one or more specific graphic sections of the image in the analytical process, and may describe those sections in ways that would not apply to the image as a whole. While this function was developed to facilitate, and may be most pertinent to, different kinds of textual transcription, as McGann has noted—it could, for example, highlight or otherwise identify, relevant text in response to a user query—it also has wider applications. In a graphic and textual multimedia object like Blake’s *Laocoön*, for example, the

inscriptions and other text could be individually mapped, identified, and transcribed, and the figures and other elements could be mapped, identified, and described. Thus, the figure on the right could be outlined and identified as Adam, and the text around his head as “Satans Wife the Goddess Nature is War & Misery & Heroism a Miser,” using SVG’s pixel-oriented `<path>` element and XML’s ability to express textual data. Such mapping makes the image granular—*this* is Adam, for example, and *this* is the inscription on nature—and less of an opaque whole; it also allows these discrete sections to be pulled out of context when necessary, when one is concerned with only one object or aspect of the image. More importantly, because SVG keeps track of pixel location in relation to the image as a whole, spatial relationships between these user-identified sections are also described and layout analysis facilitated. By analyzing the computer-identified position and orientation of the *Laocoön* inscriptions, one might be able to know the order in which they were written; and having information like the fact that Satan is to the left and Adam to the right encoded in the metadata might be useful for a certain kind of research—for example, for someone looking to see who is “sitting” at God’s right hand in such cultural objects.<sup>6</sup>

While image-based humanities computing may be inherently more user-centered than other forms of the digital humanities—because image information is often obser-

---

<sup>6</sup>More recent versions of the University of Victoria Image Markup Tool have used tags from the TEI P5 transcription module—including the `<facsimile>`, `<surface>`, and `<zone>` elements—to identify annotation areas, which has the distinct advantage—especially for textual images—of simplifying the document by keeping all markup pure TEI XML (Holmes). But the move away from SVG seems like a missed opportunity for a more sophisticated mapping mechanism. Not only can SVG more accurately describe graphic areas, its less obviously relevant elements—for example, those enabling font definition, gradients, blurs, and scripting—could play a part in any number of interpretive or deformative image-based editing activities; for example, if the image of Laocoön and his two sons were described using SVG, this markup could be quantitatively defined or compared to that of other figures of Trojan priest, or could be manipulated directly to expose the structural features of the image (McGann, “Visible and Invisible Books” 287).

vation-oriented and because searching for images plays such a huge part in the field—the annotations and connections made by creators or editors, especially in the context of the image-based edition and in how they facilitate and guide research, can be revealing and productive. Thus, the second function of Inote was to allow “one to imbed hypertext links” between editorial annotation and the digital image; creating navigational paths from text to image and vice versa, thus connecting the two basic (and different) kinds of objects” of the edition or archive (McGann, *Radiant Textuality* 95). *The Blake Archive*, for example, has linked editorial information (like their “illustration descriptions”) to each image in the archive, providing the user with one perspective of the image (especially useful when the image content or action is unclear or obtuse), but also displays links to copy information, to other plates representing the same work, and—through a few steps on the navigational path—to other objects in the archive with similar content (or “more like *this*”). Users are thus presented with a variety of ways of knowing the object or objects of interest and with—ideally—everything contained in the archive which may be of interest to the user (Fig. 3 shows how the user is presented with information in *The Blake Archive*). Because image-based editions like *The Blake Archive* are often massive and treelike, as Eaves notes in “Multimedia Body Plans: A Self-Assessment,” these paths, as well as the formation of a “matrix of information” for each object, are of the utmost importance. Image-based editions like *The Blake Archive* thus tend to focus on organization and inter-object relationships (though searching is also important), while image-based archives, on the other hand, often focus on querying and search capabilities. Thus, a group of social scientists at the University of Amsterdam is working on a project in which an image archive—in this case, a collection of art objects—is marked up and organized using a set of “controlled vocabularies,” an integrated group of pre-existing thesauri and hierarchical classification systems which have been digitized and expressed relatively

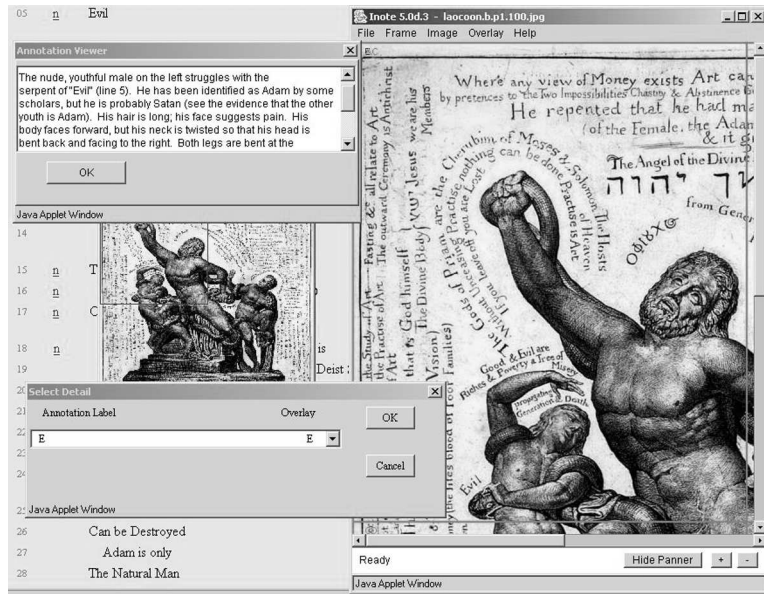


Fig. 3. “[Inote] functions most powerfully when used in conjunction with the Archives image searching capabilities, where it can open an image found by the search engine, zoomed to the quadrant of the image containing the object(s) of the search query, with the relevant textual annotation displayed in a separate window.” (“Technical Summary”).

in the RDF format. Their goal in this project has been to give “some indication of how a semantic web for images might work,” and to explore such a web’s potential for “concept search instead of pure syntactic search” and the “advanced search strategies” allowed by semantic annotation (Wielemaker et al. 150). “The Semantic Web” and “Web 2.0” are nearly meaningless buzzwords used to describe these new forms of internet metadata that “allow computers to understand” what is being processed and represented, but even this simple conception of the technology involved is useful in understanding the foundation of this project: because search terms and annotations are described in an ordered hierarchy, the computer can “understand” them, insofar as it can place them on this “tree of knowledge” or relate them to synonymous terms, and by understanding them, can offer new ways of interacting with them. The

assumption of these scholars is that by organizing “background information” in such a way, they will avoid “the search problems that currently arise on the Internet and in multimedia databases,” and that such a design will work better than the popular and often used “(semi-)structured indexing scheme that typically supports a keyword search” (147).

Following a major principle of the semantic web, the necessity of working with existing ontologies, “even if one disagrees with some of the design principles of the ontology” (Hollink et al.),<sup>7</sup> the scholars settled on four classification systems and thesauri which were “relevant for the art-image domain,” established as respected classification resources, and already existing in digital form: the Art and Architecture Thesaurus (AAT), a large thesaurus of general art terminology, used here for describing the physical object attributes (art materials or media, for example) and for some content; WordNet, “a general lexical database,” for describing action (verb) attributes of the image’s content and to fill in any gaps in the other ontologies; Iconclass, a hierarchical content-based classification system for art and iconography, for describing image content on many levels (see the slightly odd hierarchy for the action of the *Laocoön* in Fig. 4); and the Union List of Artist Names (ULAN), a biographical dictionary of 220,000 artists, for connecting the work and the artist. A derivation of the Dublin Core metadata elements is also being used for “production-related descriptors,” “physical descriptors,” and “administrative descriptors.” However, the main focus of their annotations and metadata is on pure content description, and the editors have developed a “sentence structure” template for their collected images

as a means of structuring image-subject descriptions. The subject of the image is described with the form “agent action object recipient.” Each

---

<sup>7</sup>Thus the World Wide Web Consortium notes the “the ‘ethos’ of the Semantic Web is to *share* and *reuse* as much as possible” (“W3C Semantic Web FAQ”).

statement should at least have an agent (e.g. a portrait) or an object (e.g. a still life). The terms used in the sentences are selected from the various thesauri. Multiple sentences may be used to describe a single painting. (Hollink et al.)

This scheme “avoids the problems of parsing natural language descriptions, while maintaining some of the naturalness and richness” of such descriptions. It also allows the description of the image’s setting with semantic slots for event, place, and time, to be filled with terms from the thesauri; and the editors are considering adding the Thesaurus of Geographical Names (TGN) to identify and track specific locations (Hollink et al.). Although the thesauri used sometimes overlap, the editors have added equivalence links wherever this overlapping occurs, thus verifying that no relationship is left unmarked and adding richness to these hierarchies.

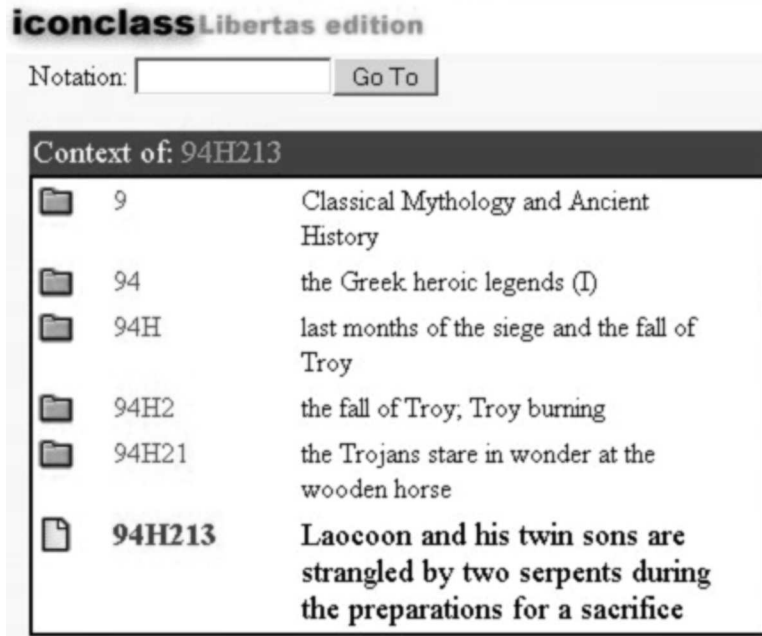


Fig. 4. The Iconclass hierarchy for the action of the *Laocoön*, as represented in the *Iconclass Libertas Browser*.

Hollink’s group is focusing on the capabilities of such high-quality image meta-



data for advanced querying, although the interpretive and editorial potentials of this metadata are equally interesting—for example, their use of WordNet to describe images has also basically laid the foundation for a kind of visual dictionary. And their markup immediately improves any type or level of image search. For example, users with even the vaguest notion of what they are looking for—say, a baroque era painting with a woman in it, or a picture of Venus (the goddess), or a picture of a woman in a bed—could easily discover Lazzarini’s painting *Venus, Cupid, and Bacchus* (Wielemaker et al. 150). Those searching for images of Venus (the goddess) would be presented with only figures of the Roman goddess; images of the planet or of the tennis player (Venus Williams), which might clutter keyword-search results, would be filtered out thanks to the disambiguation of the semantic ontologies (Wielemaker et al. 152–3). A particularly useful aspect of their search mechanism is that searches for objects that have a generic identifier—their example is the great in “great ape”—will not be flooded with results related primarily to the generic term. Using semantic ontologies, as these scholars are doing, offers distinct advantages over keyword-based systems (as incorporated in, for example, *The Blake Archive*) or other forms of image searching: having a controlled vocabulary, the ability to disambiguate search terms, and the ability to browse through related terms or over hierarchies, is a potentially huge advantage.

As Kirschenbaum notes in his introduction to image-based humanities computing, “images have not generally yielded to the same kind of regularized manipulation as electronic text, where the humanities computing community has achieved reasonable success to date in using descriptive encoding schemes to build robust textbases for structured search and retrieval,” primarily because “while various theoretical discourses have challenged distinctions between textual and visual representation, the fact remains that we manipulate images in binary formats and textual data in ASCII”

(4). While text can in some sense be managed directly, digital images must be mediated by an image viewer or editing tool; and while text can be string-searched automatically, almost all searchable data or metadata for image files—except for basic file information, or visual data about the image that the computer can understand—must be added by the creator or editor as markup. However, some image tools, including Adobe Photoshop and the GNU Image Manipulation Program (GIMP), do allow access to some forms of computer-readable visual information. Thus, using analyzing processes like color or image histograms, which identify the range or spread of color or values in a digital image in a kind of computer vision, the computer is able to recognize images “by criteria such as color, iconic shape, or by the position of elements within the image frame,” as Besser notes, and to extract such metadata from any digital image. The Blobworld project (now defunct) at the University of California, Berkeley, was developing a system for image searching using such processes, in a kind of content-based information retrieval (CBIR). Each image in their collection was processed to extract “machine-recognizable visual criteria,” and this basic image metadata was then attached to each image. The search interface allowed the user to select a search term—a “blob”—from a series of shapes based on automatically generated outlines of objects in images from the collection), and to indicate the relative importance of this “blob”—not, somewhat, or very—and the user would then repeat this process for example based on color, texture, location, and shape (“Blobworld”). The search mechanism would then retrieve images based on—or similar to—the user’s request. Though their interface required that the user be aware of an image’s general features in order to find it, the system’s accuracy was quite high, and was unique in that it needed neither background knowledge nor much in the way of editorial mediation.

On another level, image manipulation or deformation can offer visual knowledge

about “structural features” and “characteristic formal patterns” of images (McGann, “Visible and Invisible Books” 287–9). Thus the “standard filter protocols” of Photoshop or the GIMP, which distort the digital image in a variety of ways, when applied to certain digital images, may expose underlying image structures; they may, for example, reveal that many of Rossetti’s portraits of women “are dominated by patterns of interlocking vortices and spirals” (288), or may clarify “a chromatic organization” never noticed before in a familiar image (*Radiant Textuality* 84). Even “playing around with the image editor so as to twist and distort the original [image] completely” is a “useful critical [operation],” not least because people take delight in such deformation and “talk volubly about what they see” (85–6).<sup>8</sup> Such deformations “typically [reveal] not what we didn’t know but what we didn’t know we knew,” as McGann observes, and generate “what we would elsewhere and otherwise call ‘subjective’ interpretations” (86); thus leading to the idiosyncratic image description typified by Blake in the *Laocoön*. And tools like Kirschenbaum’s own Virtual Lightbox or the *Overlay* module of Kevin Kiernan’s Edition Production and Presentation Technology (EPPT), which allow users to compare, adjust, and overlay sets of digital images, may illuminate lines of influence and aspects of color use, and allow a kind of image-based “editorial palimpsest,” to use McGann’s term. Even more fascinating is the EPPT’s feature called *RamSome*, which can, for example, reconstruct missing folios (as in a manuscript) by “translating the text into image, built character by character using letters taken from [elsewhere in] the manuscript” (Iacob et al. 2). While this could be considered a deceptive editorial practice, the potential for such

---

<sup>8</sup>In such deformative activities, then, there are two forms of dialogue going on, the dialogue between the deformer and the object—exemplified in Blake’s “dialogue” with the *Laocoön* engraving, and the dialogue among those perceiving the image, which develops and describes the deformer’s own “idiosyncratic” relationship to the image.

text-to-image transmediation is intriguing, as are the possibilities here for interpretive deformation: as McGann points out, deformative scholarship often takes the form of forgery (115). Such “exceptional” tools and practices are necessary in the production of visual knowledge precisely because images—especially digital images—are such recalcitrant objects and so difficult to compare, synthesize, and analyze (Goodrum, O’Connor, and Turner 291); as McGann notes, “[unlike] language objects, once a visual object—a painting or drawing or photograph—is digitally reconstituted, it resists further moves to mirror or translate it” (“Visible and Invisible Books” 289). Thus, such tools may not only

expose characteristic formal features of pictorial works [but may also] release perception from the spell of precisely those kinds of characteristic formal patterns and open a perception of different arrangements and patterns. For the truth about works of art—textual, pictorial, auditory—is that they are, in Tufte’s word “multivariate.” (288–9).

That is, the production of visual knowledge in image-based humanities computing involves (or should involve) answering not only “What does the image mean?” but also “How do we release or expose [the image’s] possibilities of meaning?” (*Radiant Textuality* 108).

These actions or aspects of image-based humanities computing are perhaps typical of—or do something similar to—all (at least Western) forms of image editing or interpretation, historically. Thus Blake’s “editing” of the *Laocoön* engraving recapitulates these actions or aspects, although he may not always share their concern for authenticity, authority, or accuracy, nor possess the tools and mechanisms which perhaps allow more to be done with images. Blake represented his engraving, obviously, by printing it, but he also controlled how it was represented by only printing

a few copies, a meaningful act, as James points out (234). He attached some form of metadata—about the production of the work (“Drawn & Engraved by William Blake”) and about the content (in the title and inscriptions)—to the image by printing them on the same plate, thus making them inextricable. He broke the engraving down into semantic parts, as image-based humanities computing does, by “attaching” inscriptions to certain sections of the image—for example, by wrapping “Good” around one serpent, and “Evil” around the other. He controlled the progress of observation in the *Laocoön* by placing text at meaningful points in relation to the image, and by suggesting relevant external objects for further research—for example, “Where any view of Money exists Art cannot be carried on, but War only Read Matthew C X. 9 & 10.” He perhaps guaranteed that the image would be found and analyzed by mislabeling it with “keywords”—Jehovah, Satan, and Adam—that were important to his other work and mythology, though it is not known how exactly he catalogued this work. He also deformed the visual aspect of the *Laocoön* by inscribing text around it in a kind of graffiti, which perhaps emphasizes and reveals the “structural features” of the sculpture. Most importantly, though, Blake explored how to reveal the image’s “possibilities of meaning,” rather than just describing what the image means, which has been the typical focus of image-based humanities computing.

## CHAPTER III

## NEW FORMS OF IMAGE EKPHRASIS

However useful the critical approaches of image-based humanities computing may be, they are just that, critical, and however comprehensive they may seek to be, they will always only reflect a limited perspective. The semantic ontologies being used at the University of Amsterdam, for example, which promise an objective and authoritative point of view, are not necessarily the be-all, end-all of image metadata, nor are they perhaps the best way to do image ekphrasis. As Cory Doctorow, an author and editor of the popular blog *BoingBoing: A Directory of Wonderful Things* notes in his essay “Metacrap: Putting the Torch to Seven Straw-Men of the Meta-Utopia,” schemas are not neutral, “metrics influence results,” and there is always more than one way to “describe something.” That is, objectivity is not attainable even in hierarchies—many creators or editors, Doctorow points out, assume “that reasonable people, given enough time and incentive, can agree on the proper means for building a hierarchy,” but in actuality “[nothing] could be farther from the truth. Any hierarchy of ideas necessarily implies the importance of some axes over others”; artificial quantification may skew relationship descriptions; and the very nature of ontologies, or “controlled vocabularies,” may be a fatal flaw—“requiring everyone to use the same vocabulary to describe their material denudes the cognitive landscape, [and] enforces homogeneity in ideas. And that’s just not right.” The critical approaches of image-based humanities computing often limit or marginalize creative or deformative reactions to, or descriptions of, an image (the idiosyncratic approach) as exemplified by Blake’s title for the *Laocoön*—though these critical approaches are not objective, and have no sole purchase on the truth; and they limit the image’s “possibilities of meaning” by controlling how it is described or by reflecting only the analyzes or descriptions of the

few. While all metadata benefits from a level of stability and authority—else what is scholarship for?—and while administrative metadata especially requires such stability and editorial accuracy (because deformation of, or having multiple perspectives on, data-management information is typically less useful or revealing), yet if descriptive metadata especially were made open—by allowing users to describe, create search keywords for, or attach hypertext links to, the image—creative interpretations and new “navigational paths” could be supported and image queries could be improved, by reflecting this multiplicity of descriptions.

Though he has plenty to say against all metadata—reliable metadata is “a pipe-dream, founded on self-delusion, nerd hubris and hysterically inflated market opportunities”—Doctorow suggests that “observational metadata,” like folksonomy tagging, associated text or keywords developed by and for observers or users, is “far more reliable than the stuff that human beings create for the purposes of having their documents found. It cuts through the marketing bullshit, the self-delusion, and the vocabulary collisions.” And though folksonomy tagging—the “practice and method of collaboratively creating and managing tags to annotate and categorize content” (“Folksonomy”)—is equally subject to people’s lies, laziness, stupidity, and lack of self-consciousness (Doctorow), the metadata produced in such a system is open and often more accurate because it is user-oriented. Also known as “collaborative tagging,” “social indexing,” and a variety of other social web buzzwords, folksonomy tagging harnesses the energy and (often) intelligence of the usually internet masses to provide description and metadata for web pages or digital objects. Whether hosted by corporate web entities like Flickr—a site for sharing digital photographs—or user-developed and individual-centered like del.icio.us—a social bookmarking site—folksonomies typically utilize user-generated keywords and keyword collections to describe web pages, images, multimedia objects, and other web resources. Though these keywords are sus-

ceptible to inaccuracy, ambiguity, or overlapping, because they are open, yet because they exist on the surface and are easily accessible, these potential flaws are transparent, and because they are collaborative, others may correct or simply overwhelm such flaws. Even more, even this “noise”—inaccurate or potentially deformative text or keywords—is revealing and productive, by making the user think about how such text relates to the content, and by reflecting an idiosyncratic relation to the content. That is, the keyword may describe the content in a way the end-user may not have considered (but which may be revealing), and because the system is transparent, the user may look into how the creator of these keywords has used them in other contexts. Though folksonomy tagging is “not part of the underlying World Wide Web protocols” (“Folksonomy”), it is supported by an increasing number of web communities, and even Google’s search engine is a kind of folksonomy, in that web objects are retrieved and ranked based on how other web pages describe them, how many describe them, and who describes them.

The collaboration involved in folksonomy tagging works in a variety of ways. Thus, Flickr, which has an immense and energetic user base, allows users to tag their own images, and sometimes those of others, with searchable keywords, to annotate them with note boxes mapped to the image itself (similar to how the University of Victoria Image Markup Tool attaches annotations to images), and to make comments on a thread, which also are searchable. One especially interesting approach to Flickr tagging is Flickr ICONCLASSify, a self-identified “hack” that converts user-generated keywords into their Iconclass notation, thus converting these descriptive tags into a “controlled vocabulary” (Posthumus). While many scholars have argued that folksonomies, which are “bottom-up” classification schemes, and “top-down” taxonomies like Iconclass are fundamentally different, others have noted that the two are not mutually exclusive; and this system or “hack” for integrating the two forms a



“collabulary,” or a kind of taxonomy directed folksonomy (which relies on “the user interface to suggest tags from a formal taxonomy”) to facilitate higher-level organization which retains both the openness of the folksonomy and the authority of the taxonomy. The tag “bundles” used by del.icio.us, in which users may place individual tags under broader headings, are another way folksonomies integrate hierarchical organization. Another interesting twist on folksonomy tagging is Google’s Image Labeler game, which the web giant is using to index and describe almost all of the images archived by their web crawler—that is, basically every image on the internet, minus private collections and pornography. Players are paired up and shown a single image, the content of which each player must describe with a series of descriptive keywords. When the players come up with a shared term (which will then be associated with that image in future searches), they are shown another image, and the process repeats. The more specific and esoteric the agreed keyword, and the faster they go through images, the more points the players get. While image-based editors may question the value of metadata retrieved through a game, the fact is that even this basic description and these multiple perspectives are useful in image querying.

In January 2008, the Library of Congress decided to test the value of folksonomy metadata, posting more than three thousand photographs from their most popular collections, free of copyright restrictions and accompanied by their indexing information, to Flickr (Raymond). Their goal in doing so was to improve access to these photographs and to improve their own indexing information, as Matt Raymond, the Library’s Director of Communications, notes in a blog post entitled “My Friend Flickr: A Match Made in Photo Heaven” (along with his sneaking suspicions about the project’s possible failure):

If all goes according to plan, the project will help address at least two

major challenges: how to ensure better and better access to our collections, and how to ensure that we have the best possible information about those collections for the benefit of researchers and posterity [...] One of our goals, frankly, is to learn as much as we can about [the power of user communities] simply through the process of making constructive use of it.

Each of these photographs was made open for image annotation, tagging, and comments for any Flickr user, in the hope that these users would contribute, for example, “missing key caption information such as where the photo was taken and who is pictured,” which could potentially enhance the quality of the bibliographic records for the images” (Raymond). If the project were a success, Raymond suggests, the Library of Congress would make more of its more than “14 million prints, photographs, and other visual materials” available on its Flickr photostream. And their experiment has apparently been a huge success, both in widening their audience and in generating metadata for their images. Raymond posted again about the project two days later, noting that after just one day, there were about “1.1 million total views on our account,” which is a lot, even on Flickr. Most of these images have now been viewed individually over fifty thousand times, and most of these viewers, from all over the world, would never have seen the photographs otherwise. Even more interesting, though, is the user-supplied tagging and annotation which has accumulated for these photographs. One photograph, David Bransby’s *Woman Aircraft Worker* [...], described in the Library of Congress classification with just five subjects—*Vega Aircraft Corporation, Airplane Industry, Women—Employment, World War, 1939–1945, and United States—California—Burbank*, has more than twenty annotations, seventy searchable tags, and eighty comments (see Fig. 5 for a screenshot of the photograph in its Flickr context). Although many of the annotations are sim-

ply inane—for example, “The hairstyle is just nice”—others, though casual in tone, raise important question: for example, whether the photograph was staged or depicts an actual worker—“but no working human would be dressed like her!!!!FAKED PHOTO???” and “Given that she is in Burbank, I’m surprised she is wearing long sleeves in June.” Even more, the tags generated by the Flickr community are extensive and varied, and obviously useful. They range from simple description—for example, *red, wires, gloves*—to interpretation—*rosie the riveter, war effort, working glamour*—to production information—*David Bransby, slide film, 4x5*. Though there is no way for the community to rank or weight these somewhat disorganized tags (which might be useful), they are even so far more descriptive and useful (for both understanding and finding the image) than the Library of Congress classification.

But is there a place for deformative—or even random or absurd—description in such folksonomy tagging? Would Blake have been banned from using Flickr? What if one were to tag the photograph described above as *Marilyn Monroe*, for example, or even *ice cream*? Such tags would no doubt increase the exposure of the image by misleading users searching for these terms—perhaps a worthy effect in itself if the image is something needing to be shared—and some of these users might even enjoy being misled, or consider the appearance of such a result fortuitous.<sup>1</sup> More importantly, if someone was searching for Marilyn Monroe, and this image was one of the results, both Monroe’s blond hair (in contrast to this unnamed woman’s black) and her famous red lipstick (in similarity to this woman’s bright lipstick) would be emphasized, and the user would be led to compare and contrast their conception of Monroe, filtered or mediated through the other search results, with the woman in this photograph. But the real question is whether these keyword tags reflect any

---

<sup>1</sup>Misleading description and hypertext links are a fascinating internet meme, and one with a long history; for more, look up “rickrolling” on Wikipedia.

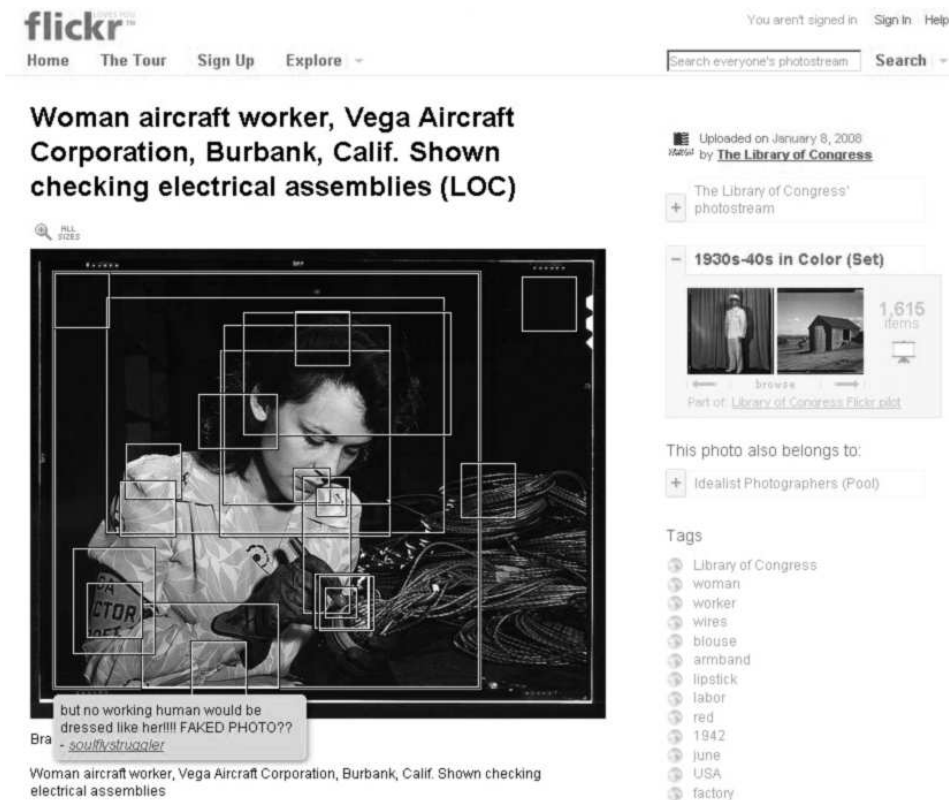


Fig. 5. Screenshot of *Woman Aircraft Worker* [...], with annotations, from Flickr.com.

of the image’s “possibilities of meaning.” By tagging this photograph as *Marilyn Monroe*, the deformer might be calling attention to the woman’s lipstick, might be calling the user to think about how this woman is like or unlike Monroe, or might be setting the stage for some sort of elaborate revisionist history. *Ice cream* might be a slightly less revealing tag, as the image hardly sustains such a description, although, if the end-user were to analyze how the deformer used this tag in other contexts, for other images, an interesting set of idiosyncratically related images, connected by some common theme in the mind of the deformer, might be discovered. As noted above, it is often the context involved in, or the dialogue created by, deformance that makes it a critically interesting activity (McGann, *Radiant Textuality* 84–86). On the other

hand, the negative results of such a tag, by decreasing the precision of the search, might outweigh any interesting or revealing effects; and if no context is provided nor reasons given for such inaccurate tags, they might be largely worthless.

And if image description is to be opened to the internet community (certainly a disparate set of persons), it might as well be opened to the machines as well. Pattern recognition for images—that is, computer-generated content description—is as yet, at least for the humanities scholar and the typical computer user, in its infancy. While pattern recognition for textual images is fairly advanced and open—Optical Character Recognition (OCR) is an extremely powerful technology for converting images of text (especially printed text) into ASCII format—there is no parallel for iconographic and artistic images; and while pattern recognition technology for digital imagery is being developed by biometrics companies, little of their research is available to or accessible by the humanities scholar. But if background information and descriptive metadata were related digitally to the content-based information retrieval practices and procedures of the Blobworld project in a diverse and large image collection, image-based humanities computing might have at least a primitive version of pattern recognition, capable of suggesting content description for processed images. And the deformance almost certainly involved in such a primitive system would even be revealing: if such a system were to describe the dress of the unnamed woman in the photograph as wallpaper, for example, or flowers, based on the dress’ color, texture, location, and shape (“Blobworld”), even those ultimately inaccurate description reveal real and different aspects of the dress.

Opening descriptive metadata to a community (or to *the* community), and perhaps to the machines, is not only necessary for effecting the requisite multifaceted and “multivariate” description for images, such a system would help to “automate” descriptive markup, by sharing or distributing the work of description. The admis-

sion of the University of Amsterdam editors that “our approach relies to some extent on manual annotation” (Wielemaker et al. 154) is almost certainly an understatement, and even the most basic descriptive markup for images often requires hours upon hours and level upon level of elaborate, hand-coded markup. While there are tools that facilitate the actual formation of (XML) markup—the interface developed by the University of Amsterdam editors, for example, or the University of Victoria Image Markup Tool—the aspect of image-based humanities computing that yields most readily to automation (or semi-automation) is this image description. Fortunately, image description is often the most time-consuming part of this practice; unfortunately, it is also the most important, and by decentering editorial mediation, automation could devalue this descriptive metadata. But fears about authority, accuracy, and authenticity seem overwhelmed by the potential of opening up descriptive metadata. Images are of interest to a number and variety of disciplines and persons, and the “question of what is needed [in descriptive metadata] for image retrieval [and knowledge] is inherently multidisciplinary” (Jørgensen 312). Iconclass, for example, which identifies allusions and classical figures in images, is an excellent “controlled vocabulary” for art historians, but is less obviously useful for bibliographers or those studying aesthetics. There is no way creators or editors can fully anticipate every way their images will be used or searched for. Opening description to these persons and communities allows different aspects or objects to be identified, different terminology to be used, and different mediations to be performed.

And while folksonomy tagging as done in Flickr might currently be the best method for collecting image metadata, because such systems depend more on users and are relatively easy to implement, it might not be the best way: the editing of this metadata is not generally sourced to the community, for example, and higher level organization does not come naturally in such tagging. The ideal might be

a model more along the lines of Wikipedia: a massive, ever-growing and evolving metadata collection open to additions and editing by the community and to a variety of terminologies in description. Such a model might seem to lack the stability of “controlled vocabularies” or static metadata, because the description could change almost constantly, but this perceived instability is unfounded, as wikis can produce a unique and static URL for any stage of the article or page; and it might lack the aggregation of metadata of Flickr’s tagging (where the more seems to be the merrier), as the article might be pared down almost constantly. But such a system would allow users to describe images in whatever way they see fit—whether with deformative descriptions or relatively objective ones—to remove descriptions that are inappropriate or useless, and to defend and revert to their descriptions if they were to be needlessly removed. Such a system could merge ontologies and integrate both “top-down” taxonomy and “bottom-up” folksonomy, combining the best of both worlds in an effort to provide a truly comprehensive description for an image. By existing outside of any one project or archive, this descriptive metadata could be used as a semantic resource by any image-based project or archive (whether in the humanities or elsewhere), each of which might choose to use some different part of the metadata. If copyright allowed, the image could be represented as reference as well, but if not, the description alone would be valuable; even more, other resources referencing or representing the image might be linked, in addition to the description, creating a rich “matrix” of knowledge for the image.

## CHAPTER IV

## NEW FORMS OF TEXT ILLUMINATION

On the flipside of the textual description of images—image ekphrasis—is text illumination or illustration, or “textual iconography,” as Eduardo Urbina has titled the practice. While, for any given illustration there may be any number of factors at work—patronage or marketing considerations, for example, play a huge role in what gets illustrated, as do changing ideas about the role of illustrations (and of books), about art in general and the image-as-such, and numerous social, political, religious, or economic concerns—and any number of relationships to be considered—for example, those between illustrations by the same artist or creator, where ideas or objects may be reused, and between illustrations of the same work, where influence or the “anxiety of influence” (cf. Harold Bloom) may be at work—but typically the most important and interesting relationship is that between the illustration and the text behind it. In their essay, “Visual Knowledge: Textual Iconography of the *Quixote*, a Hypertextual Archive,” Urbina et al. (working from McGann’s earlier divisions) define three categories of such text to illustration relationships: works “in which the textual and visual components were designed together and must be considered as a whole” (which could also apply in the image to text, or image ekphrasis, direction), works in which “illustrations provide the author’s interpretation of his [or her] own writings” (for example, the “illustrated novels of Dickens or Thackeray”), and works in which “the visual component is a representation and/or interpretation of the text by artists unknown to the author of the text” typified by the “textual iconography” of the *Quixote* (Urbina et al. 3). To which we could add a fourth category, exemplified by Blake’s *Laocoön*: works in which a pre-existing image is pulled out of context and applied to a textual work to clarify or complicate that text. Each of these kinds of



works may relate text and illustration differently, but generally both the form and the content of the image are connected to both the form and the content of the text.

The analysis of how the form and content of the illustration is related to the form and content of the text—the question of what gets illustrated, how, and why—is one of the most fascinating aspects of the study of “textual iconography.” For example, why do some texts get illustrated more, and more richly, than others? Is it simply the verbal imagery, or perhaps the illustrator’s fancy, that determines what gets illustrated and how? The responses of readers—and of illustrators—to a text are always different: different scenes will catch the illustrator’s attention, and new illustrations will be made based on personal interest. But why do certain action, objects, and scenes recur more often than others? These questions get even more interesting when one focuses on the illustrations of one work over time. Which takes precedence: how the work has been illustrated before, or the artist’s own reading of and desire for the work? Tracking what exactly gets illustrated seems the best way of answering these questions, and high quality descriptive image metadata for a collection of illustrations could be especially revealing. Thus, an analysis of the descriptive metadata for a collection of all the illustrations of one illustrator might identify the favorite actions, figures, and scenes, or other preoccupations, of the illustrator, where they appeared, and how this creator or artist interpreted both the text and his or her role as an illustrator. Such an analysis for a collection of the illustrations of a single textual work might show what scenes or textual passages were most popular and to whom historically, and how the function of illustration has changed: in the case of Chaucer’s *Troilus and Criseyde*, for example, it might show that the most popular scenes for illustration historically are those of Troilus falling in love with Criseyde in the temple, of the meeting of the two in the house of Deiphebus, and of Troilus delivering Criseyde to Diomedes; that beds and walls are oft-depicted objects; that the

number, as well as the quality, of illustrations both decreases and increases over time, as literacy improves—and lessens the need for illustrations—and the cost of making books declines, and that the nadir in regard to both number and quality is Wynkyn de Worde’s 1517 edition with its five generic—though fascinating—woodcuts.

And if some form of descriptive metadata could be extracted from and related to discrete sections of the illustrated text, the exact nature of an illustration’s relation to the text might be revealed. Of course, not all illustrations necessarily have a textual counterpart (for example, illustrations which depict the author in the act of writing), but they often do, and text is often inherently transmedial. To this end, the editors of the Mnemosyne website have experimented with using Iconclass, the image content classification system, to describe texts, as described in this unnecessarily defensive and abrupt section of their essay “Using Iconclass for the Iconographic Indexing of Emblems”:

If you have already drawn the conclusion that our approach to iconography is extremely narrow-minded because we seem to be using Iconclass for just about any task that comes along, then I might as well tell you we have even experimented with indexing the mottoes, epigrams and commentaries with it. You will already consider us as lost cases anyway.

Using an image classification system like Iconclass to describe text is certainly a strange idea, and many texts will not yield to such classification—there is not an Iconclass equivalent for every linguistic object or textual event, and there can thus be no continuity of markup using Iconclass. At the same time, such an image-based point of view can identify one very important aspect of text, its visually-oriented, transmedial nature. For the authors of this essay, indexing textual works with Iconclass no doubt helped to clarify where emblem illustrations were derived from in the

text, where emblem texts—and images—were related to each other, and where there were lines of influence among these texts and images. Of course, descriptive image metadata in the form of visually-oriented keywords (as used in Flickr, for example) could also be related to the text, but using Iconclass is actually a smart strategy: because it is a “controlled vocabulary,” and because all of its terms have image-based counterparts, one can simply scan a text for instances of Iconclass terms to develop such textual descriptive metadata. Thus one might expect to find a higher number of Iconclass instances (described content objects) in discrete textual sections that have proven popular with illustrators historically—for example, in the three or four stanzas describing Troilus falling in love with Criseyde in the temple—and one would be correct.

But such descriptive metadata for text and textual works has another effect, one almost certainly unintended by the editors at *Mnemosyne*: all of a sudden, these marked up texts are suggesting illustrations relevant to their content, assuming that this specific image-based metadata has a physical counterpart in some existing image object (a relatively safe assumption considering the diversity of images in existence). That is, if *Troilus and Criseyde* were marked up using Iconclass, for example, and if there were a massive collection of images also described using Iconclass, then the Iconclass notations for the three or four stanzas depicting Troilus falling in love—*prince*, alphanumerically identified as 44B15121, *Troilus*, as 95A(TROILUS), *temple*, as 12E61, *crowd*, as 46A7, *love at first sight*, as 33C211, and so on—would match up with the Iconclass notations for Eric Gill’s illustration of the scene, for example, or, more interestingly, with perhaps an illustration from an edition of Boccaccio’s *Filostrato*. Such illustrations would “work,” or not, depending on the quality and extent of the metadata involved, but, as Blake has shown, an image only needs to relate to the text on a few key points (the engraving of *Laocoön* chosen perhaps

primarily to illustrate a struggle) for it to be revealing. That is, even if such metadata were deficient, the slightly inappropriate illustrations suggested would, by deforming the meaning of the text slightly, be fascinating and educational, and might reveal a lot about how illustrations are meant to work—for example, by identifying why a suggested illustration would be inappropriate if created by a book illustrator, or by identifying minor visual aspects of a text—in addition to functioning at least as some kind of postmodern set of illustrations in themselves.

Here is how such a system for self-illustrating texts, or automated transmediation, might work. First, a textual work would be regularized and broken down using a digital natural language processing (NLP) tool like MontyLingua—an open source text processor to which one can feed “raw English text” and which can extract “subject/verb/object tuples, [...] adjectives, noun phrases and verb phrases, [...] people’s names, places, events, dates and times,” and other semantic information, and which is also able to resolve contractions and strip inflectional endings. This broken down and regularized text would then be scanned against Iconclass’ vast keyword database, and matches would be marked in and attached to the text, or the creator or editor could use background information to identify which discrete textual sections were visually-oriented and what keywords would apply. These keywords, and especially those of keyword clusters (which would suggest an “event” needing illustration), would then be entered, in different combinations if necessary (the database might not contain a match for every visual aspect of a textual section), into the search engine of a richly described image database like Flickr. A ranking mechanism or algorithm could be implemented, so that the system could “learn” what is involved in a “good” illustration, based on whether the creator or editor agrees that any suggested image works as an illustration for a textual event and using machine-readable content-based information and keyword clusters to keep track of the ones that work.

Such a mechanism could also be taught by the creator or editor identifying already existing illustrations—for example, ones that exist in a printed book—of a certain text as appropriate illustrations for that text. In any case, though there would be many false positives—at least at first, and especially if there were no filtering—there would also be a number of relevant and novel illustrations. These images could then be attached or related to a text in an electronic edition or elsewhere—barring copyright complication—whether to simply show what is going on in the text, to reveal a happenstance or fortuitous interpretation of the text, or to complicate the text further.

A short, object-oriented prose text in modern English—Ernest Hemingway’s “A Very Short Story”—has been selected to illustrate this illustration process. Because the mechanisms used here are very primitive and require a lot of manual processing, brevity is essential—hence “A Very Short Story” and not just a short story—as is the use of modern English, which can be scanned and searched much more easily (especially in Flickr). As with many of Hemingway’s works, this story is visually descriptive—especially considering the word to description ratio—words are carefully chosen, the author never uses an adjective where a noun would work, and so on. Briefly, it tells the story of an unnamed wounded soldier and a nurse, Luz, who fall in love while he is in a hospital in Padua. They try to get married, but do not succeed, soon after which he leaves—they have a quarrel at the train station—with Luz planning to come to America when he can support her. They correspond, until Luz sleeps with a major and decides to break off their relationship. The story closes with him contracting gonorrhoea from a sales girl “riding in a taxicab through Lincoln Park.”

Such a brief summary is perhaps a typical critical response to the narrative, but it neglects almost everything that makes this story visually interesting, suggesting

that the visual aspects of the text are non-essential (traditionally) and that visual knowledge of a text works in acritical ways. The first paragraph of the story, broken down, reveals a number of image-oriented words and text phrases: *evening, Padua, roof, top of the town, chimney swifts*, and so on. Entering all of these terms as keywords on Flickr produces no results; the archive is limited and there is no digital photograph available that contains all of these objects or aspects. But when one limits the search to a few crucial (and sometimes generic) keywords, valuable and relevant results appear. Thus the keywords *roof, man, woman, and birds*, searched for in Flickr's Creative Commons content—that is, a collection of images licensed through Creative Commons which can be freely used (typically) if one credits the creator or artist—produce some fascinating results. *The Place to Be*, a digital photograph by Flickr user Auer1816, depicts a man and a woman standing slightly apart on a covered deck (or top of a building) with birds on the roof, looking out over a body of water in the dusk (probably), and offers perhaps the best fit to this scene in Hemingway's story (see Fig. 6). Few places in Padua have such a view of the sea, the man is standing (the story suggests he could not at this point), and the birds are at rest rather than “in the sky,” but the photograph represents many essential features of the textual scene. Even more, this image also suggests meanings only potentially available in the text: the two figures are slightly too far from each other, perhaps indicating their final separation—the text neither supports nor denies this relative position—and the tone of the photograph is certainly pensive or wistful, which the text may or may not be at this point. Even the deformed aspects of the image—the man's pose, the birds, and the presence of the sea—have revelatory meaning: the sea could symbolize eternity, for example, belying their love, the birds' rest could prefigure the quiet atmosphere of the upcoming paragraphs, and the man's stance might reflect how he feels despite his injury.



Fig. 6. “they carried him up onto the roof [...] There were chimney swifts in the sky.”

Potential illustrations for the rest of the story are also suggested. Thus the keywords *quiet*, *church*, and *prayer*, derived from and for the third paragraph, bring up illustrations like *a prayer in St. Germain*, by user *chelseagirl* (see Fig. 7): not especially interesting or revealing, but it works as a simple illustration and it certainly reflects the tone or mood of the paragraph. An illustration for the fifth paragraph, *goodbye* by G.L.o. (see Fig. 8), in which the two main characters say goodbye at the train station in Milan, is more interesting. In the image, two figures face each other in bidding a melancholy farewell, but the train is distinctly contemporary. This would no doubt be a common occurrence in using a typically photography-centered database like Flickr (and is perhaps one reason Flickr would not be the best image database to use in such a system); but though this anachronism deforms the text, it also perhaps makes it more relevant, or at least suggests that the story is timeless. Finally, Flickr user *litherland*’s photograph *Kristen, Eliot* (Fig. 9), another slightly



Fig. 7. “they went into the Duomo and prayed. It was dim and quiet [...]”

generic illustration for the conclusion of the story, keeps the soldier’s final actions tastefully out of sight, though the figure passing the taxicab is perhaps fittingly mysterious, or ominous. Even more interesting is the fact that, though it is not tagged as such, Flickr user litherland uses this image to discuss the Eliot Spitzer “escort scandal” (in the comments): another level of meaning added to the image which partially coincides with the meaning added by Hemingway’s text.

As a comprehensive set of illustrations for Hemingway’s story, these images per-





Fig. 8. “When they had to say goodbye, in the station at Milan [...]”

haps do not quite work: tone, features, color, and other elements disrupt the flow between images across the story—though this disconnect perhaps emphasizes and reflects the non-equivalence and non-serial nature of all textual work. This lack of flow, felt at least by the reader, reflects the novelty of such a mode of illumination—one based solely on the text’s visual content and content-based keywords, not the story, style, marketing, and so on—though it also recalls Wynkyn de Worde’s innovative use of generic woodcuts in illustrating his texts in the sixteenth century, where, for example, the printer used a woodcut of a man being treated by a doctor plus one of a noblewoman identified by a flying banner as “Cresyde” to depict Criseyde visiting Troilus in bed in the house of Deiphebus. Individually, however these images can “pull out” hidden meanings—or potential meanings—from the text, through the ways they slightly or not so slightly deform the nature of the text. Such deformance can enhance reading by forcing the reader to search the image and text for incongruity; and by transmediating the text into an image, something “resistant” and paradoxi-



Fig. 9. “while riding in a taxicab through Lincoln Park [...]”

cally open to interpretive deformance (McGann, “Visible and Invisible Books” 289), it allows new ways of accessing images. As metadata for image collections continues to expand and improve, the possibilities and potentials of such a system of automated transmediation will increase both for scholarly communities and for others—for example, creators or artists wanting to illustrate a textual work using only relevant Early Modern woodcuts or only blue-colored images, pedagogues seeking to illustrate a text simply to assist reading comprehension, or literary scholars simply seeking new ways to approach a text, will have more resources to do so.

## CHAPTER V

## SUMMARY

Images—whether multimedia works like the *Laocoön*, manuscript facsimiles, art objects, or the digital photographs of Flickr—have two primary functions in the humanities: they are, perhaps most importantly, objects to be analyzed, described, and interpreted, but they are also the subject, or the mediation, for themselves or for other kinds of information. They are things to be broken down and understood, a critical action that requires the mediation of analysis, description, and interpretation, but they are also more than the sum of the (textual) description of their parts. It seems obvious that images share (or transmit) much of their information (or essence) visually, and that the image itself is “so detailed that one could never describe it [textually] with sufficient accuracy” (Flanders 308), and thus requires visual representation or self-mediation. Images may also function as the mediation for other, often textual materials, in illustrations, for example, or in the “visualizations” of textual data (Kirschenbaum, “Editor’s Introduction” 3), offering new ways of knowing these materials (through clarification or complication). Of course, these graphic mediations (for themselves and for other objects) are also open to the critical approach, but, in moving beyond textual understanding, they suggest that such an approach may not be enough. Associating an image with a text, as illustration does, is a creative action, as well as a potentially critical one, and the textual description of an image, because images are so complex and because they have nearly limitless “possibilities of meaning,” can and should be creative as well, to expose the multifaceted nature of the image.

Image-based humanities computing, which “brings advanced visual and visualization tools to bear on the objects and artifacts of cultural heritage” and which

involves building “editions, collections, or self-identified ‘archives’ that place digital images at the center of their interface (and database)” (Kirschenbaum, “Editor’s Introduction” 3), has typically focused on the critical, object-oriented approach to images: on representing images, attaching metadata to them, breaking them down into discrete semantic parts so that they can be known and used in a variety of ways, linking editorial annotations and images—or images and other images—to facilitate or guide research, identifying all relevant images for all kinds of user queries, or using the computer as computer to practice some form of interpretation—or in some cases deformance—upon images. These actions or aspects of image-based humanities computing are perhaps typical of—or do something similar to—all (at least Western) forms of image editing or image interpretation, though with perhaps more sophisticated tools and mechanisms to do so, especially now: as Goodrum, O’Connor, and Turner note (in 1999), “[recent] advances in image storage capacity, network bandwidth, processing power, and display resolution for digital images of all kinds have enabled development of new tools and reconfiguring of trusted old tools in the last few years” (291).

Such a critical approach is important and is no doubt the beginning of some kind of image-based knowledge, but at the same time image-based humanities computing seems to lack definition, seems unnecessarily limited in its approach, and seems unaware of what can be done with the technologies involved. The kinds of knowledge images offer, as well as the kinds of images of interest and the variety of approaches allowed by images, often goes unexamined, and the textual mediation required in making images usable or knowable often goes unacknowledged. Image-based humanities computing is clearly yet paradoxically textually oriented: not only is such description and analysis—as well as the computer markup underlying the field—generally performed in and through text, but bringing images “into the fold of structured data”

means making them more like “searchable [or machine-readable] text” (Flanders 303, 308). And many of those engaged in image-based humanities computing fail to realize the inherently deformative nature of critical representation, description, and analysis, as well as the critical possibilities of embracing and integrating this deformation. As Jerome McGann notes in “Visible and Invisible Books: Hermetic Images in N-Dimensional Space,” all critical representation “deforms its object,” because “the object perpetually shifts and mutates under the influence of its perceivers,” and involves only “a certain perspective on the object,” one which is perceptive insofar as it understands its “own powers and limitations” (287). Even more, description and analysis provide only a certain perspective, not only because the image contains too much information to be fully described (cf. Flanders 308), but also because different people will see different things in an image. And while poor description is often due to people’s lies, laziness, stupidity, or lack of self-consciousness, as Cory Doctorow notes in “Metacrap: Putting the Torch to Seven Straw-Men of the Meta-Utopia,” misalignments between image and description can be valuable (especially if intelligently designed). If one were to change or deform the obvious meaning of an image in a description, that description might be considered inaccurate, but it might also reflect one revealing perspective on the image. By analyzing such inaccurate description in the context of the image, trying to make sense of how this text describes this image, the user would be making productive connections in non-traditional, atypical ways, by analyzing semantic aspects of the image that could work in more ways (or have more meanings) than one. Images are multifaceted and, to use Edward Tufte’s term, “multivariate” (qtd. in McGann: 289), and can sustain—and may need, to reveal all of their potential aspects—a variety of descriptions or identifications, including those which may be deformative or typically (traditionally) exclusive. However useful the critical approaches of image-based humanities computing may be, they are just that,

critical, and however comprehensive they may seek to be, they will always only reflect a limited perspective; and the tools and technologies of this field—the image classification schemes and archives, for example, or the image manipulation protocols—can do much more than simply assist the critical approach.

In entitling the *Laocoön* plate “[Jah] & his two Sons [...]” and filling it with descriptive text, Blake shares the focus of image-based humanities computing on images as things to be broken down, described, and understood. But Blake’s classification and description, deformative in misreading the image, reveals the true, deformative nature of such mediation and the need for a more open system, one which accepts the idiosyncratic response, and which allows observers to record how they interpret an image. This openness in developing descriptive metadata is perhaps best accomplished digitally through semantic web or social web technologies like folksonomy tagging; and the experiment performed by the Library of Congress of uploading more than three thousand of its most popular images to Flickr to facilitate access and to improve its image metadata by allowing user tagging and description of these images (Raymond, “My Friend Flickr”), seems to have been a huge success. But while folksonomy tagging as done in Flickr might currently be the best method for collecting image metadata, the ideal might be a model more along the lines of Wikipedia: a massive, ever-growing and evolving metadata collection open to additions and editing by the community. At the same time, Blake’s act of pulling a pre-existing image out of context and applying it to a new textual work suggests a new function for images and the highly structured image databases of the field, to clarify or complicate any textual work through graphic mediation; and the system for automated transmediation outlined above shows the new and creative potential for the classification strategies and highly structured image databases of the field, and for images as graphic mediation.

## WORKS CITED

- Auer1816. *The Place to Be*. 2008. *Flickr*. 25 Mar. 2008 <<http://flickr.com/photos/brianaauer/2295529406>>.
- Besser, Howard. *Introduction to Imaging*. Ed. Sally Hubbard and Deborah Lenert. Rev. ed. *The Getty*. 24 Mar. 2008 <[http://www.getty.edu/research/conducting\\_research/standards/introimages](http://www.getty.edu/research/conducting_research/standards/introimages)>.
- Blake, William. *Laocoön*, copy B. *The William Blake Archive*. Ed. Morris, Eaves, Robert N. Essick, and Joseph Viscomi. 15 Apr. 2008 <<http://www.blakearchive.org>>.
- Bransby, David. *Woman Aircraft Worker, Vega Aircraft Corporation, Burbank, Calif. Shown Checking Electrical Assemblies*. 1942. Library of Congress. *Flickr*. 25 Mar. 2008 <[http://www.flickr.com/photos/library\\_of\\_congress/2179930812/in/set-72157603671370361](http://www.flickr.com/photos/library_of_congress/2179930812/in/set-72157603671370361)>.
- Chaucer, Geoffrey. *The Noble and Amerous Au[n]cyent Hystory of Troylus and Cresyde, in the Time of the Syege of Troye. Co[m]pyled by Geffraye Chaucer*. London: Wynkyn de Worde, 1517. *Early English Books Online*. 21 Apr. 2008 <<http://eebo.chadwyck.com>>.
- chelseagirl. *a prayer in St. Germain*. 2006. *Flickr*. 25 Mar. 2008 <<http://flickr.com/photos/chelseagirlphotos/229424658>>.
- Doctorow, Cory. "Metacrap: Putting the Torch to Seven Straw-Men of the Meta-Utopia." 2001. *The Well*. 12 Mar. 2008 <<http://www.well.com/~doctorow/metacrap.htm>>.

- Eaves, Morris. "Multimedia Body Plans: A Self-Assessment." *Electronic Textual Editing*. Ed. Lou Bernard, Katherine O'Brien O'Keeffe, and John Unsworth. New York: Modern Language Association of America, 2006. 210–23.
- Eaves, Morris, Robert N. Essick, and Joseph Viscomi, Eds. "Copy Information." *Laocoön*, copy B. *The William Blake Archive*. 15 Apr. 2008 <<http://www.blakearchive.org>>.
- . "Illustration Description." *Laocoön*, copy B. *The William Blake Archive*. 15 Apr. 2008 <<http://www.blakearchive.org>>.
- . "Technical Summary." *Laocoön*, copy B. *The William Blake Archive*. 15 Apr. 2008 <<http://www.blakearchive.org>>.
- . "Textual Transcription." *The William Blake Archive*. 15 Apr. 2008 <<http://www.blakearchive.org>>.
- Edwards, Jane A. "The Transcription of Discourse." *The Handbook of Discourse Analysis*. Ed. Deborah Schiffrin, Deborah Tannen, and Heidi E. Hamilton. Cambridge: Blackwell Publishers, 2001. 321–48.
- Flanders, Julia. "Trusting the Electronic Edition." *Computers and the Humanities* 31.4 (1998): 301–310. *SpringerLink*. 18 Feb. 2008 <<http://www.springerlink.com>>.
- "Folksonomy." *Wikipedia, the Free Encyclopedia*. Wikimedia Foundation, Inc. 2 Mar. 2008 <<http://en.wikipedia.org/w/index.php?title=Folksonomy&oldid=194696329>>.
- G.L.o. *goodbye*. 2006. *Flickr*. 25 Mar. 2008 <<http://flickr.com/photos/gasperoni/309048361>>.



- Goodrum, A. A., B. C. O'Connor, and J. M. Turner. "Introduction to Special Topic Issue of *Computers and the Humanities*: Digital Images." *Computers and the Humanities* 33 (1999): 291–2. *SpringerLink*. 18 Feb. 2008 <<http://www.springerlink.com>>.
- Hemingway, Ernest. "A Very Short Story." Malaspina University-College. 10 Apr. 2008 <<http://www.mala.bc.ca/~lanes/english/hemngway/vershort.htm>>.
- Hollink, Laura, Guus Schreiber, Jan Wielemaker, and Bob Wielinga. "Semantic Annotation of Image Collections." Vrije Universiteit. 18 Feb. 2008 <<http://www.cs.vu.nl/~guus/papers/Hollink03b.pdf>>.
- Holmes, Martin. *The UVic Image Markup Tool Project*. TAPoR@UVic. 18 Feb. 2008 <[http://www.tapor.uvic.ca/~mholmes/image\\_markup/index.php](http://www.tapor.uvic.ca/~mholmes/image_markup/index.php)>.
- Iacob, Ionut, Kevin Kiernan, and Alex Dekhtyar. "Edition Production Technology: an Eclipse-Based Platform for Building Image-Based Electronic Editions." *EPPT*. 18 Feb. 2008 <<http://eppt.org/~emil/publications/ach05-ept.pdf>>.
- Iconclass Illustrated Libertas Edition*. 2008. Rijksbureau voor Kunsthistorische Documentatie. 27 Mar. 2008 <<http://www.iconclass.nl/libertas>>.
- James, David. "Blake's *Laocoön*: A Degree Zero of Literary Production." *PMLA* 98.2 (1983): 226–236. *JSTOR*. 18 Feb. 2008 <<http://www.jstor.org>>.
- Jørgensen, Corinne. "Access to Pictorial Material: A Review of Current Research and Future Prospects." *Computers and the Humanities* 33.4 (1999): 293–318. *SpringerLink*. 18 Feb. 2008 <<http://www.springerlink.com>>.
- Kirschenbaum, Matthew. "Editor's Introduction: Image-based Humanities Computing." *Computers and the Humanities* 36.1 (2002): 3–6. *SpringerLink*. 18

- Feb. 2008 <<http://www.springerlink.com>>.
- . “Refining Our Notions of What (Digital) Images Really Are.” *Institute for Advanced Technology in the Humanities*. 18 Feb. 2008 <<http://iath.virginia.edu/ach-allc.99/proceedings/kirschenbaum.html>>.
- litherland. *Kristen, Eliot*. 2008. *Flickr*. 25 Mar. 2008 <<http://flickr.com/photos/litherland/2333599431>>.
- Liu, Hugo. “montylingua [...]” 2004. *MIT*. 24 Feb. 2008 <<http://web.media.mit.edu/~hugo/montylingua>>.
- McGann, Jerome. *Radiant Textuality*. New York: Palgrave, 2001.
- . “Visible and Invisible Books: Hermetic Images in N-Dimensional Space.” *New Literary History* 32.2 (2001): 283–300. *Project MUSE*. 15 Apr. 2008 <<http://muse.jhu.edu>>.
- Pinelli, Bartolomeo. [Plate 4]. Illustration. *Le Azione Piu Celebrate del Famoso Cavaliere Errante Don Chisciotte della Mancia*. By Miguel de Cervantes. Roma: R. Gentilucci, 1833. Texas A & M University. *Cervantes Iconography Project*. Ed. Eduardo Urbina. 18 Mar. 2008 <<http://www.csdl.tamu.edu:8080/dqiDisplayInterface/displayMidImage.jsp?edition=764&image=1833-Roma-Gentilucci-01-007.jpg>>.
- Porter, Dorothy. “Examples of Images in Text Editing.” 2007. *Digital Humanities*. 18 Feb. 2008 <[http://www.digitalhumanities.org/dh2007/abstracts/paper\\_250-porter.pdf](http://www.digitalhumanities.org/dh2007/abstracts/paper_250-porter.pdf)>.
- Posthumus, Etienne. “ICONCLASS.” 2007. *Iconclass*. <<http://www.iconclass.org/flickr>>.

- Raymond, Matt. "My Friend Flickr: A Match Made in Photo Heaven." 16 Jan. 2008. *Library of Congress Blog*. 29 Feb. 2008 <<http://www.loc.gov/blog/?p=233>>.
- . "Flickr Followup." 18 Jan. 2008. *Library of Congress Blog*. 29 Feb. 2008 <<http://www.loc.gov/blog/?p=237>>.
- Schwartz, Joan. "Negotiating the Visual Turn: New Perspectives on Images and Archives." *American Archivist* 67.1 (2004): 107–22. *Society of American Archivists*. 18 Apr. 2008 <<http://archivists.metapress.com>>.
- Smith, Martha Nell. "Computing: What Has American Literary Study To Do with IT?" *American Literature* 74.4 (2002): 833–57. *Project MUSE*. 18 Mar. 2008 <<http://muse.jhu.edu>>.
- Sundt, Christine. "Digital Projects Past and Present—Survivors or Fossils?" University of Oregon. 18 Feb. 2008 <<http://darkwing.uoregon.edu/~csundt/chart2004.doc>>.
- Urbina, Eduardo, Richard Furuta, Stevan Escar Smith, Neal Audenaert, Jie Deng, and Carlos Monroy. "Visual Knowledge: Textual Iconography of the *Quixote*, a Hypertextual Archive." *Literary and Linguistic Computing* 21.2 (2006): 247–258. *Oxford Journals*. 18 Mar. 2008 <<http://llc.oxfordjournals.org>>.
- "Using Iconclass for the Iconographic Indexing of Emblems." 2005. *Mnemosyne*. 18 Mar. 2008 <<http://mnemosyne.org/about/projects/emblems>>.
- Wielemaker, J., A. Th. Schreiber, and B. J. Wielinga. "Supporting Semantic Image Annotation and Search." *Annotation for the Semantic Web*. Ed. Siegfried

Handschuh and Steffen Staub. Tokyo: IOS Press, 2003. 147-55. *NetLibrary*. 18 Feb. 2008 <<http://www.netlibrary.com>>.

“W3C Semantic Web FAQ.” 2008. Ed. Ivan Herman. *World Wide Web Consortium*. 18 Mar. 2008 <<http://www.w3.org/RDF/FAQ>>.

## VITA

Name: Jacob Hohmann Brown  
Address: Mailstop 4227 TAMU  
College Station, TX 77843  
Email Address: kardeiz@gmail.com  
Education: B.A., University of Dallas, 2005  
M.A., Texas A&M University, 2008

The typist for this thesis was Jacob Brown.