2016

Textual Curation

Krista Kennedy
Syracuse University

Follow this and additional works at: https://surface.syr.edu/wp

Part of the Digital Humanities Commons

Recommended Citation

This Article is brought to you for free and open access by the College of Arts and Sciences at SURFACE. It has been accepted for inclusion in Writing Program by an authorized administrator of SURFACE. For more information, please contact surface@syr.edu.
Textual Curation

Krista Kennedy
The Writing Program, Syracuse University

krista01@syr.edu
239 Huntington Beard Crouse Hall
Syracuse, New York 13066

Vitae
Krista Kennedy is an Assistant Professor of Writing and Rhetoric in the Writing Program at Syracuse University, where she studies and teaches digital rhetoric, authorship, rhetoric of technology, and professional writing. Her single- and collaboratively-authored work has appeared in College English as well as Rhetoric and the Digital Humanities, the second edition of A Guide to Composition Pedagogies, and other edited collections. With Rebecca Moore Howard, she recently edited the College English special issue on “Western Cultures of Intellectual Property.”
Textual Curation

Abstract

This article explores textual curation as a conceptualization of authorship and composition within large information structures that is heavily based on the cannon of arrangement. This work is often undertaken through distributed collaboration, thus complicating traditional conceptions of authorial attribution and agency. Central curatorial processes include critical recomposition of prior texts along with the development of small and often invisible textual elements such as architecture, metadata, and strategic links. I offer a grounded definition of textual curation that draws from traditional curatorial fields such as Museum Studies and Library studies as well as Writing Studies’ own subfield of Technical Communication, which focuses heavily on recomposed, collaboratively produced texts. Selected Wikipedia articles serve as case studies for examining live curatorial work in open, collaborative environments.

Keywords.

Authorship, textual curation, digital composition, collaboration, information architecture, metadata, tagging, linking
“I wanted to teach students how to write the small texts that make the world turn,” a mentor said to me one autumn afternoon years ago when I asked how she had come to teach writing. In the years since that conversation, the question has come up again and again in conversations with other writing teachers, and we always come back to the small things. We want to help students learn to make a good claim, develop effective written arguments, a solid rhetorical analysis, a well-told narrative. Each capable visual analysis, each accessibly designed web page, each competent podcast helps our students make the world turn. The same is true in even smaller textual units: a rhetorically effective tweet, a carefully styled sentence, a well-crafted phrase, an ethical citation.

We’re comfortable with such small textual units, that is, as long as they’re readily identifiable as essayistic writing. We want words, and quite a few of them—at least 140 characters, not only a word or two. Consequently, we tend to reserve for the technical communication classroom—or not teach at all—the ofteninvisible compositional skills that transform an informative website that contains words and images into a useable, accessible, unified text with searchable content. These strategic links, recomposed texts, metadata elements, and information architectures are suasive elements that contribute heavily to the ethos of digital arguments, and they help both writers and readers realize the full potential of digital environments.

In the years since Kathleen Blake Yancey’s call for increased pedagogical attention to multimodal composition in 2004’s “Made Not only in Words: Composition in a New Key” and later in the NCTE report Writing in the 21st Century, writing in digital contexts and in social media has become an increasingly integral part of composition pedagogy. As a field, we’ve moved closer to naturalizing an understanding of writing
more akin to the one that Andrea Lunsford proposed in her keynote address to the 2005 Computers and Writing conference:

Writing: A technology for creating conceptual frameworks and creating, sustaining, and performing lines of thought within those frameworks, drawing from and expanding on existing conventions and genres, utilizing signs and symbols, incorporating materials drawn from multiple sources, and taking advantage of the resources of a full range of media. (171, emphasis mine)

Since then, the generic landscape of the writing classroom has changed. Teachers and students at a wide range of institutions have taken up many forms of blogging, screencasting, podcasting, and video development, as well as emergent short genres such as tweets. However, as our attention to the many modes and mediums in which writing happens has grown, we have remained primarily concerned with recognizable essayistic writing, whether it be in a status update, a Vine, or a collaboratively written website. In these contexts, a carefully composed 140-character missive or a six second video can comprise a narrative argument. Even in assignments that are rich with generative discussions of ethical borrowing and remix, the composition of textual or visual communication frequently remains the sole object of assessment in digital composition classrooms. This focus is necessarily accompanied by a concentration on familiar genres that facilitate assessment of argumentation, narrative structure, style, and other textual facets that are quite similar to those of a more typical essay. Indeed, this is one of the common arguments deployed against naysayers when it comes to digital pedagogies: these digital assignments teach these same essential compositional and rhetorical aspects
as print texts, but in dynamic digital environments that better prepare students for current work environments and acknowledge emergent communication practices.

As vital as these developments have been to our research, to our classrooms, and to our fulfillment of university missions, they do not yet usually offer a broad focus that includes the small and often invisible writing skills that are vital for managing both personal information ecologies and larger digital texts. This is partly due to the simple fact that good information architectures, carefully constructed taxonomies, and usable navigation elements such as metadata and strategic links are largely unnoticeable when done well, as is filtered information that has been recomposed with more attention to rhetorical effectiveness than to demonstrating the sort of original authorship most often valued in university writing. But it is also because we don’t often really consider these forms of composition to be fully formed, essential writing skills that are essential elements of courses in digital writing.

Jodie Nicotra (2009) has powerfully addressed elements of this issue, arguing that understanding findability and information architecture as vital elements of composition is fundamental to teaching forms of digital literacy that account for the social, networked nature of contemporary writing. The discipline of Rhetoric and Composition has long been interested in rhetorical contexts of literacy, but as she argues, our comfortable ways of teaching essayistic writing no longer fully account for the ways that writing functions on the web:

Now more than ever the focus is much more on the organization of the total network than on the individual producer of texts. … the sheer amount of information with which we’re dealing now and the medium in which it primarily
occurs has perhaps given the importance of organization over individual authorship a heightened intensity. Thus, the issue of findability [and information structures] is an important one for contemporary rhetoric and composition. (266)

She focuses on folksonomic tagging in social sites like Flickr as an example of small, collaboratively produced texts that negotiate and facilitate persuasive findability. This distributed development of information taxonomies is both deeply rhetorical and indisputably categorized as writing, Nicotra contends, creating a space for students to develop grounded understandings of audience awareness and a metacritical awareness of network participation.

Teaching the full range of literacies necessary to develop a functional website that goes beyond a simple collection of smart textual and visual narratives requires helping students understand this product as part of the living ecology of humans and objects that make up the web. This move requires a fundamental reconsideration of which units and types of text constitute writing. Doing so, writes Johndan Johnson-Eilola (2005) in his discussion of emergent reading processes and skills, requires us to learn to understand learning and work in new ways: Creativity is no longer the production of original texts, but the ability to gather, filter, rearrange, and construct new texts... Users need to have available to them as much information as possible and robust methods for moving information around, virtually and physically (134, emphasis mine).

This contention remains accurate and well-stated, but several years on, users are not only learning new methods for reading interfaces, but also new methods for writing within those interfaces. We can no longer simply teach students how to rhetorically compose
texts and embed images in digital environments without also asking them to become familiar with the labor of curating that work: the basics of project management, information gathering and filtering, strategic linking, metadata management, and basic site architecture.

I suggest that these essential processes and the conceptualization of authorship and composition that informs them constitute a category of compositional craft called *textual curation*. The arrangement-driven composing and small texts necessary for this work are not less-important or diminished writing, but carry with them the value attached to writing-as-skilled-craft. Given its frequently collaborative nature, it requires becoming comfortable with forms of authorial agency that are distributed and socially situated. Nicotra argues that in the case of folksonomies, “rhetorical agency and intention become much more complicated, because invention is revealed as not simply the product of an individual, isolated mind, but as a distributed process driven by the interaction of a multitude of users” (273). Textual curators must always contend with distributed collaborative environments and, consequently, distributed agency because they are always and ever working to arrange prior texts into innovative, flexible textual ecologies. Recomposition processes, which I discuss below, require working with a multitude of previously published texts that will be filtered and recomposed into a new unified text. This repurposing of others’ writing pushes against traditional expectations of original writing that inform our grading and tenure guidelines. The same is true for other curatorial tasks that require us to reconsider the requirements of original, single-authored writing, such as writing metadata for another’s article¹, collaborative editing, large-group

¹ For extensive discussion on collaborative development of metadata, see Ball, 2013.
architecture development, and the like.

The stages of composition required in the curatorial process, which include filtration, recomposition, and designing structures for findability and navigation, also demand closer attention to the canon of arrangement. Successful curation requires a broader conceptualization of textual organization that moves outside of individual texts, such as the individual encyclopedia articles I examine in this piece, and into the ecologies they exist within (Brooke, 2009), both in their original state and in the new ecology they form through recomposition. Addressing this need means teaching the curatorial moves required when writers are managing digital information in both personal and business contexts, whether it be categories on an organizational blog, a basic business website, a citation library on Zotero, tagged images on Flickr for a nonprofit, a Storified archive of tweets collected via hashtags at a convention, or the information architecture of a digital humanities project. Regardless of the purposes for which students deploy these environments, they must negotiate the social nature of digital writing, which also includes not just collaborative processes but also creating audience-centered texts that readers can navigate intuitively.

Teaching curation builds on the field’s long tradition of teaching collaborative writing and offers a number of potential complications that help prepare students for writing on the open Web. Learning to undertake curation in complex digital contexts helps students negotiate composition in live environments that includes contributors they do not know and who encounter the text in multiple temporalities. They may collaborate after the fact with authors of the texts they are filtering and recomposing; asynchronously or in real time with contemporaneous community members who manipulate their text by
not just adding words but manipulating it through code; and then with still other contributors who may alter the text years in the future. Because of the sheer mass of information that must be managed, recomposed, and arranged in order to successfully develop a large digital text, curators are frequently working in collaboration with prior contributors over extended periods of time, as seen in Wikipedia’s development.

Collaboration occurs recursively throughout production stages as well as after the fact, when the readers begin to write their own experience of non-narrative texts by finding their own paths through it via cross-indexing, links, or individual choices. The textual curator’s contributions, then, come through honed research skills, astute arrangement, clear recomposition, and curatorial information management.

This sort of collaboration can be understood as dialogic, following Lisa Ede and Lunsford’s landmark 1990 study of collaborative writing. They understand dialogic collaboration as a loosely structured, fluid collaborative practice in which “one person may occupy multiple and shifting roles as a project progresses” (133). The intensely dialogic nature of curatorial collaboration proceeds in exactly this way, especially in large collaborative projects where collaborators may wear multiple hats in a single work session, let alone over the life of the project. A curator may add primary text to a page, then edit another person’s contributions to another text, audit and clarify metadata, then switch to conducting research in order to gather broad information for filtration and recomposition. In another session, they may make sure that navigational links are up to date. All the while, other contributors intersect with the first curator’s work, both asynchronously and in real time. The work may occur with little direct communication,

---

2 For more on the reader-as-writer of non-narrative texts, see Eco, Magli and Otis, 1989.
discussion may happen behind the scenes, or arguments may break out and then move to arbitration. The internet’s inherent affordances of speed and reach, which Laura Gurak (2001) has written about at length, further enhance both the collaborative potential in digital environments and the potential for fluid shifts in curatorial roles. These affordances have fostered the increased deployment of distributed collaboration in both organizations (Spinuzzi, 2007) and individual business practices (Pigg, 2014). Similarly, it has given rise to what Yochai Benkler (2006) termed commons-based peer production, the sort of distributed collaboration found in projects like Wikipedia that depend on many contributors from all over the world working through a non-hierarchical collaborative process develop projects that are frequently based on creative works that are in the public domain. While it is not typically possible to recreate this sort of broadly distributed environment in the composition classroom, it is entirely possible to develop whole-class collaborative assignments that ask students to work in a distributed manner with a group that is significantly larger the small groups that we so often rely upon.3

A number of helpful terms for understanding the collaborative development of larger information structures have emerged from Writing Studies scholars’ studies on workplace writing, which offers a natural arena for examining this sort of work (Dias, Freedman, Medway, & Par, 1999; Reyman, 2001). Scott Jones (2005) describes common workplace writing tasks such as “document borrowing,” which includes single-sourcing and adapting prior documents for new contexts. Through these and other collaborative

---

3 For extensive discussion of wiki-based pedagogies that employ commons-based peer production, see Cummings, 2009; for more on specific aspects of using wikis in the composition classroom, see the essays in Cummings and Barton, 2008; for more on assessing wikis in the technical communication classroom, see Barton and Heiman, 2011.
tasks, all done as work-for-hire, the writers in his study transitioned from “being the sole creators of mainly independent units of texts to being among the creators of highly interconnected parts of a very large unit of text.” Other scholars have described this process of re-composing and structuring texts in a variety of ways: “layering” (Geisler, 2001), “environment selecting and structuring” (Prior & Shipka, 2003), and “textual coordination” (Slattery, 2005); the resulting assemblages of texts are understood as “genre ecologies” by Aviva Freedman and Graham Smart (1997) as well as Clay Spinuzzi and Mark Zachary (2000). Both Jones and Shaun Slattery suggest that these changes are at least partly due to changes in technology, which encourages broader collaboration by making the tasks of coordination more efficient.

While textual curation shares some aspects in common with each of these terms, it differs fundamentally when examined in context: the wilds of socially accepted knowledge rather than controlled corporate environments. Each of the above studies examined the work of re-composition within workplace contexts. While textual curation can certainly be performed within institutional contexts, it is also common in forms of knowledge work that are commonly practiced outside of corporations, such as fan-written gaming guides (Luce, 2014), community-curated media archives (Lewis, 2013), forum discussions devoted to various hobbies, and perhaps the most prominent example, Wikipedia. It is also not necessarily driven by digital technology, although it is most commonly practiced in these environments today; the foundational English encyclopedist Ephraim Chambers discussed these same issues in the 1728 preface to the Chambers’ *Cyclopædia*^4^. For the purposes of this article, though, I confine my discussion to the

---

^4 For more on Chambers, see Yeo (2001) and Kennedy, Lamoine & Révauger (2013).
digital environments that we and our students encounter daily and offer a case study of curatorial found in Wikipedia, one of the most commonly accessed reference works available today.

As the most prominent contemporary digital encyclopedia, Wikipedia illuminates the process of curating large information ecologies in significant ways. Small contributions by more than 45,000 contributors have resulted in nearly 4 million articles in the English edition alone as of this writing. The fundamental affordances of wiki platforms make them well suited for curatorial tasks, especially in projects that rely on distributed labor. Wikis are networked, collaborative, anonymous, and dialogic systems that rely on central database structures accessed by multiple users. The users themselves move within a network of intertextual and interpersonal connections, particularly in sub-communities such as those that have developed among the broader, sprawling community of Wikipedians. Wiki platforms enable what has become known as radical collaboration “because of [their] total freedom, ease of access and use, simple and uniform navigational conventions, and apparent lack of formal structure” (Leuf & Cunningham, 2001, 16.) Wikipedia’s policies of openness are evident in throughout its interface, not just in the “Edit” tab that enables anyone to contribute, but in the “View History” tab that displays all edits to the page since its creation. Consequently, it provides the largest and richest database of curatorial work in the world that documents what is now the longest experiment in open-access text creation. Studying its archives offers an opportunity for a long view of curatorial processes.
Wikipedia is frequently described as an encyclopedia that anyone can edit, and in its early days this was unequivocally true: while founder Jimmy Wales and co-founder Larry Sanger were overseers of the site, all contributors were considered more or less equal, and contributions were immediately visible on the Web. This principle of equality became part of the community parlance: all contributors were (and continue to be) referred to as “editors” with no distinction made between various curatorial or writerly activities. I adopt this terminology except in my references to curators and similarly do not make distinctions between writing and editing simply because these distinctions do not exist as hard and fast roles within Wikipedia. Any editor may move between a variety of contribution types within a single work session. Because of this multi-level, structural freedom, the development community must constantly negotiate not just content and editing issues, but issues concerning architecture, taxonomies, metadata, and the undergirding philosophies of the project (Zachry & Morgan, 2010). In other words, the central tasks of curation are always being negotiated within the community, even though a central set of tenets known as the Five Pillars provides a philosophical basis for this work.\(^5\)

Despite the hundreds of thousands of registered editors, only a handful significantly impact the project. Over the years, the estimation has fluctuated: in 2005,

\(^5\) Wikipedia’s Five Pillars outline the fundamental principles by which it operates:

- Wikipedia is an encyclopedia.
- Wikipedia in written from a neutral point of view.
- Wikipedia is free content that anyone can edit, use, modify, and distribute.
- Editors should treat each other with respect and civility. (http://en.wikipedia.org/wiki/Wikipedia:Five_pillars)
Wales estimated that “2% of the users do 75% of the work,” while more recent empirical studies have suggested even more startlingly that one-tenth of 1% of editors—approximately 4,200 people—contribute nearly half of the site’s content (Priedhorsky et al, 2007). This dedication demonstrated by relatively few editors is what renders the distributed work of composing Wikipedia curation rather than random composition by individuals dropping in (seemingly out of the digital skies), making changes, and then never returning. This cohesive, central group of editors has both an understanding and commitment to the goals of Wikipedia and works to make it a curated encyclopedic project rather than a loose collection of links and opinions. Over time, the community has come to adopt cultural definitions of curation and build affordances for the work of curation into its own technological structure.

**Vernacular and Interdisciplinary Definitions of Curation**

The word curation has been in the vernacular since the mid-seventeenth century, but it’s enjoyed an explosion over the past five years. It’s a term I first began using one morning in 2007 as I struggled to describe my research on the labor of composing reference texts composition to my writing partner, who was himself working on rhetorical aspects of museums. These days, curation has moved out of the museum and into popular discussions of working with almost any everyday collection, most particularly digital ones. By 2009, the South by Southwest Interactive conference featured a panel on “Curating the Crowdsourced World” (Bekman, Johnson, McEvoy, Hostetler & Trapani, 2009) and the term “community curated work” began to appear as an alternative to the older term “user-generated content” in discussions about wikis.
(“Community Curated Works,” 2010). In April of that year, the Silicon Valley Insider claimed that “curation is the new role of media professionals ... Separating the wheat from the chaff, assigning editorial weight, and—most importantly—giving folks who don’t want to spend their lives looking for an editorial needle in a haystack a high-quality collection of content that is contextual and coherent.” (Rosenbaum, 2009).

These days the word has proliferated even further, and is usually meant to describe the curator’s primary task as one of filtering an immense amount of information through the critical lens of one’s own sensibilities. (Most often, one’s aesthetic sensibilities.) There are curation apps such as PearlTrees and Storify to deploy, curation communities like Pinterest to join, curation contests to enter in online home design and personal fashion communities. *The New York Times* noted that curation is now the provenance of most any creative type: “among designers, disc jockeys, club promoters, bloggers and thrift-store owners, curate is code for “I have a discerning eye and great taste” (Williams, 2009). Everyone is a curator, it seems, as we struggle to make meaning within the information overload of the postmodern world, most particularly the unyielding data stream that is the web.

The problem with this increasing ubiquity of the term is that along the way, we’ve robbed it of its meaning. Describing and assigning meaning to curation as mere filtration, aggregation, or collection strips this compositional work of the craft performed through the curator’s labor. Those of us who work in fields that have adopted “curation” to describe filtered and recomposed compositions may forget that curation is a specialized craft and field of study in multiple curation-focused disciplines that award advanced degrees, train specialists, launch distinguished careers, and create robust scholarship.
I draw definitions primarily from Museum Studies and Library Studies because of their rather obvious connections with carefully-arranged large texts. Both work to order collected, filtered knowledge in ways that are publicly accessible not only textually but also in terms of findability and structure. The curatorial skills involved are hardly new: Patrick J. Boylan, Professor Emeritus of Heritage Policy and Management at City University London, traces the advent of specialized training in museum curation to the École du Louvre’s 1870 offerings, and notes a proliferation of museum studies degree courses in South America and the UK during the 1920s and 1930s (2006, 424). Library Science has been concerned with findability, metadata, and circulation since Melvil Dewey launched the field’s first formal program in the U.S. at Columbia University in 1887. Both disciplines, along with Art History, understand the skill and labor entailed in curation in specifically defined ways that are closely tied to learned craftsmanship. Their definitions can and should shape our own understanding of textual curation, since they shed light on specialized forms of compositional work: most particularly, the work of composing texts that are comprised of collected, filtered, ordered information.

This work, which I have sketched in broad strokes above and explore in greater detail later, bears significant resemblance to Boylan’s description of eighteenth-century museum curation:

Scholar-curators undertook almost all of the museums’ specialized work: acquiring collections, specimens, and works of art, researching cataloguing, and documenting their collections, and interpreting and communicating their significance through the museum’s permanent display galleries, temporary
exhibitions, publications, and educational programs such as lectures and guided visits. (2006, 418.)

To say that a text has a curator conveys an appropriately heavier emphasis on the specific performance of authorial agency demonstrated in critical assessment, re-composition, and arrangement of previously disseminated work, moving the emphasis further from individual originality. It also more accurately describes the unending work of curating a living text and/or body of knowledge that is in constant flux, much as the work of curating a museum is never complete until the museum is shuttered. Johnson-Eilola has written about the ways that symbolic-analytic tasks such as filtering, sorting, connecting, synthesizing, and sharing have become central for scholars who seek to efficiently work with ever-growing streams of information (2010, 48). As University of South Australia Library and Information Science professor Ross Harvey asserts, curation of library collections is concerned with actively managing data for as long as it continues to be of scholarly, scientific, research, administrative, and/or personal interest, with the aims of supporting reproducibility, reuse of, and adding value to that data, managing it from its point of creation until it is determined not to be useful, and ensuring its long-term accessibility, preservation, authenticity, and integrity. (2010, 8)

This work of collaboratively collecting, filtering, recomposing, taxonomizing, and managing information is essential not only to museum and library curation but also to textual curation. So too are the “invisible texts” that function in conjunction with the primary reference text, such as controlled vocabularies, taxonomies, strategic linking and
cross-indexing, and metadata. Arranging, interconnecting, and recomposing are essential processes for curatorial writing; the relevant skills may manifest themselves in the everyday life of a writer through something as simple as a stack of books relevant to a chapter that is under construction or as complex as a cross-referenced, tagged, categorized reading blog that also links to external sources. Such structuring is deeply familiar work to any writer who works extensively with large amounts of digital information.

This understanding of the labor processes associated with composing in the current environment leads us to a different conceptualization of collaboratively produced digital information structures, whatever their genre: open, interconnected, and not necessarily finished in the ways that we previously deemed projects to be complete once they were published and distributed as print artifacts. We typically understand print compositions as reaching a terminal point in development, whether it be an essay, a collage, a film, or a symphony. Print encyclopedias are always instantly dated, and for the past 300 years in the West, we have always understood that a new edition is forthcoming. Digital environments change this: as multiple commentators have noted, Wikipedia’s value lies in its constant updates, ongoing project development, and minimal distribution costs. We understand it as always and ever under construction. This distinction is important if we are to understand the always-in-process nature of textual curation, particularly in networked environments. They are process, not product, as Shirky has famously discussed, and require terminology that acknowledges as much (2008, 119). James Purdy casts this open-ended process as a valued aspect for study by scholars of composition, writing “Wikipedia (and wikis more generally) asks us to reexamine our
expectations for the stability of research materials and who should participate in public knowledge making. … Wikis and other Web 2.0 technologies make more visible the complex, rich, messy, processes usually kept behind the closed doors of the academy. [By allowing users to make changes, Wikipedia contests] the notion that online texts, even encyclopedic or scholarly texts, are ever permanently finished” (2009, W352-356). Accounting for this open-endedness more explicitly in our pragmatic approaches to writing and pedagogy requires making space for generative messiness in our understanding of ongoing curatorial compositions.

The authors of the UCLA Digital Humanities Manifesto 2.0 describe curation as simultaneously persuasive, technical, and technological, focused on making arguments through objects as well as words, images, and sounds. ... It is a medium with its own distinctive language, skill sets, and complexities; a medium currently in a phase of transformation and expansion as virtual galleries, learning environments, and worlds become important features of the scholarly landscape. Curation also implies custodial responsibilities with respect to the remains of the past as well as interpretive, meaning-making responsibilities with respect to the present and future (2009).

Curation, then, is a rhetorical, dynamic skillset that pays close attention to the techniques required for “taking advantage of the resources of a full range of media” – the very skills that Lunsford called for a decade ago. The essential aspects of curation are among those found in her definition of writing itself: creating conceptual frameworks through the process of building taxonomies and architectures, expanding on existing conventions and genres, utilizing signs in symbols in the form of navigational aspects and strategic
linking, and recomposing text by incorporating materials drawn from multiple resources. Curation is writing, regardless of how small or invisible its texts might be.

Wikipedians have also developed their own cultural understanding of curation that has become naturalized within the community, particularly over the past few years. The tasks of textual curation, which are both collaborative and rhetorical, provide a view into what the daily negotiation of articles can involve and, more specifically, precisely what forms of labor involved in managing and improving the text. As with any textual genre, encyclopedias require the writing tasks that we most often study and teach in rhetoric and composition classrooms: textual composition and revision alongside careful design and integration of visual elements. As Purdy points out in his examination of Wikipedian writing processes and composition theory, writing in Wikipedia takes the tenets of composition public (2009). In wiki pages, which are always and ever in open-ended draft stages, editors with higher levels of involvement leave tags that point to certain tasks that would strengthen the page. Many of these tasks entail what the community calls wikification, or “formatting according to Wikipedia style” (Ayers, Matthews & Yates, 2008, 205). Wikifying can include converting plain text to hypertext, adding interwiki links, or restructuring the articles into the usual article format of an overview followed by subsections, as well as the specific tasks mentioned above.

Until recently Wikipedians described this work as “editing,” in keeping with their discourse convention of everyone being an “editor.” Over the past two years, though, wikification has been increasingly described by the community as curation. Recent
investments by the Wikipedia Foundation, which supports Wikipedia and is responsible for much of the funding and infrastructure that drives technological improvement within the site, has resulted in changes that rhetorically conceptualize curation within the project culture. For example, the Foundation has driven a number of efforts to make workflows more streamlined for dedicated editors⁶. In 2012, the Foundation supported the development of the Page Curation tool suite, which provides experienced editors with a feed of new pages and a curation toolbar that facilitates more efficient review of articles. The Curation Toolbar also formalized the term curation as part of the community’s nomenclature.

⁶ For years, editors relied on a constantly updated Wikipedia page containing a list of new pages and third-party apps such as Twinkle and Huggle to monitor new edits and vandalism.
This Javascript panel “enables editors to retrieve basic information on the page’s creation and status, mark a page as reviewed, tag it, mark it for deletion, send WikiLove to page creators—or jump to the next page on the list” (“Wikipedia:Page_Curation,” 2012). Many of the tasks that the Curation Toolbar renders explicit are those that have been prevalent in the system for many years: basic editing, metadata development, sourcing, interwiki linking, categorization, monitoring compliance with the neutrality policy, and monitoring vandalism. This tool explicitly articulates these tasks as central to the labor of maintaining a rigorously constructed digital encyclopedia, and at the same time it facilitates this labor and reinforces the essentialness of these tasks. By embracing this terminology and including tasks typically identified as “editing” (such as “copy edit” and “writing style”) under the broader heading of curation, it defines textual curation as a tasks that encompasses editing but extends beyond it to include a richer range of labor types.
badge on their User page. By publicly identifying themselves as Curators, they lay claim to a broader skill set that includes editing but also extends beyond it to the tasks of wikification and quality control. Curatorship, then, also becomes a self-selected identification that carries with it affiliation, social capital, and ramifications within the community.

With the goal of more closely examining the sort of iterative work hinted at by the development of the Curation Toolbar, I selected a limited sample of communities and articles for study. While this sort of limited sample is not broad enough to reveal generalizable information in the same way that an automated study of thousands of articles would be, it makes use of informed judgment to closely examine data. The results provide a grounded and very specific look at the development processes of standard Wikipedia articles. Like the majority of Wikipedia articles, these communities and articles focus on non-controversial topics that have not recently been in the news or seen particularly heavy traffic. They are not hot vandalism targets, as articles on timely political issues are, nor have they been selected for Featured Article status on the front page of Wikipedia and thus seen a community push to improve them. They are simply basic, informative articles that have been steadily built out by interested members of the community over more than a decade.

The process of curation is often recursive, and evolutionary metaphors do not apply. Instead, due to constant contributions and the unending nature of the project, we’ll see filtration practiced in the midst of recomposition, and wikification necessarily happening throughout the process. The technological
affordances of the environment contribute to this recursiveness and the varied
symbolic nature of a single activity, since a link to internal or external source can
simultaneously function as verification of valid recomposition, cross-indexing,
and taxonomic navigation within the system.

Navigating Curatorial Communities and Projects

Filtration happens in two important ways in a large curated project like
Wikipedia: taxonomization of information, which also drives the self-selected sorting of
editors. The project’s information architecture helps both readers and editors navigate the
sprawling structure. Given this article’s focus on production, I begin by examining the
ways that it assists contributors in finding the best topical areas for them to contribute
their expertise and the communities that facilitate development of these areas. As
Priedhorsky, et al’s study indicates, the contributors who stay with the project long-term
tend to be invested and committed; unsurprisingly, strong communities form around
common curatorial projects and interests. Thousands of WikiProjects curate topics in
diverse umbrella categories such as “games and toys,” “arts and culture,” and
geographically specific umbrellas such as “Africa” or “the Americas.” Subcategories
further direct contributors to more specific areas of interest, so that a Lego enthusiast
drilling down from “games and toys” may choose to work specifically on the Lego
project, which focuses on all Lego-related articles on Wikipedia, but not on the Bionicles
Taskforce, which is categorized as a related but entirely different concern. These fine
distinctions in taxonomy are just one element that both enables the curatorial work as
well as demonstrates the editors’ commitment to careful, considered functionality.
Many smaller communities form around curatorial tasks, as the Military History group has. This group’s page announces that its members are “dedicated to improving Wikipedia’s coverage of topics related to military history” and openly invites participation in improving any of the more than 100,000 articles within the scope of the project. This sort of directed curatorial effort is not unusual in the system. Hundreds of other WikiProjects curate topics in diverse umbrella categories such as “games and toys,” “arts and culture,” and geographically specific umbrellas like “Africa” or “the Americas.” Subcategories further direct contributors according to more specific areas of interest, so that a Lego enthusiast drilling down from “games and toys” may choose to work specifically on the Lego project, which boasts 122 members and focuses on all Lego articles in Wikipedia.

For subject-area communities, the job of producing relevant texts serves as a central impetus. As a result, many sub-communities form around interest-based curatorial tasks, as the Military History Group has. This group’s community page announces that its members are “dedicated to improving Wikipedia’s coverage of topics related to military history” and openly invites participation in reaching any of its current four major milestones: having 750 of its articles approved for featured article status (77.7% complete), achieving 500 items of featured content other than articles (80.2% complete), improving 2000 of its articles to “good” status (90.8% complete) and having at least 10% of its articles rated B-Class or better (79% complete) (“WikiProject:Military_History,” 2013). Each of the classifications is linked in order to better direct members; for instance, by clicking
on “featured articles” the reader is taken to an alphabetized list of articles on military history that need improvement.

As is typical of most groups, the group directs active project management by maintaining an Open Tasks page listing articles that need general or specific improvements. Many of these improvements focus on wikification: 35,972 articles need work on reference or citation, for instance. Others need work on “coverage or accuracy,” “structure,” and “photographs.” At this writing, the assessment list was backlogged, as was the list of articles that needed project tags fixed. The page also lists content currently under various sorts of review: peer reviews, “good” article candidates, featured article candidates, and the like. These task lists highlight the community’s focus on curatorial tasks as essential to community standards of good composition and point community editors toward these tasks as priority items. This sort of broad yet detailed project management functions as the topmost level of curatorial work: defining the parameters of a project and identifying the labor to be done.

**Incremental Filtering and Recomposition**

The majority of *Wikipedia* articles begin life as “stubs,” or very brief summaries of only a few words that are posted as a way of creating a page that others can add to. Ayers, et al describe stubs as “incomplete—by definition, they lack something vital—but they are often useful and well written. Approximately 70 percent of *Wikipedia* articles are still classified as stubs” (2008, 7). Unlike print, the iterative potential of wikis allows multiple editors to make small,
incremental changes to a text that add up over the life of the article. Since these incomplete pages provide spaces for continued expansion, they’re viewed by the community as open invitations for contribution, especially by casual contributors who may happen upon the page via a Google search. This sort of dialogic yet asynchronous collaborative development is a common aspect of curatorial practice.

Let us take as an example the incremental growth of five articles that as of this writing are a decade or more old—quite old in terms of Internet time: the articles on falconry, fortification, garden, minerals, and trigonometry. During their development cycles, all of them have all grown incrementally and exponentially, as Table 1 shows.

<table>
<thead>
<tr>
<th>Article</th>
<th>Creation Date</th>
<th>Creation Length</th>
<th>Current Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Falconry</td>
<td>1/16/03</td>
<td>21</td>
<td>8,177</td>
</tr>
<tr>
<td>Fortification</td>
<td>3/31/03</td>
<td>32</td>
<td>5,562</td>
</tr>
<tr>
<td>Garden</td>
<td>2/27/02</td>
<td>1</td>
<td>2,487</td>
</tr>
<tr>
<td>Minerals</td>
<td>12/12/01</td>
<td>236</td>
<td>12,563</td>
</tr>
<tr>
<td>Trigonometry</td>
<td>9/8/01</td>
<td>157 (69)</td>
<td>2,403</td>
</tr>
</tbody>
</table>

This selection is not entirely random. They are part of a book-length comparative historical study of curation in Wikipedia and the 1728 Chambers’ Cyclopaedia. Article topics are pulled from Chambers’ taxonomy of knowledge (Chambers, ii) and further refined by the necessity for topics to remain relatively stable in meaning over the nearly 300 years separating these primary artifacts.
Table 1: Shows article growth in word count since creation. The Trigonometry article was immediately edited for plagiarism after its creation, reducing its affective start length to 69 words.

Only rarely has text growth been due to additions of large chunks of text. Rather, they are the result of thousands of small edits: analysis of the Trigonometry article included more than 3,552 edits,\(^8\) and other edit histories are of similar length. This sedimentary accretion of words and facts illustrates a central philosophy of *Wikipedia*: eventualism, “the idea that things will eventually improve if you leave them around long enough” (ibid, 349). When early detractors argued that article stubs were worthless and that errors immediately rendered the project suspect, Wikipedians argued in turn that the project was designed for long-term growth, and that slowly articles would grow and errors be corrected. The artifacts analyzed here certainly did grow incrementally\(^9\). The first creation of the Garden article began with a single word on a page: Garden. That brief article was quickly deleted by another editor who thought it was a possible duplicate of the Gardening page, but the creator swiftly clarified their intention of creating a distinct page. Both the Minerals and Trigonometry articles began with more substantial word counts, although the Trigonometry stub was immediately flagged for plagiarism and the offending second paragraph was removed, bringing the stub start length to 69 at the second edit.

\(^8\) It’s possible to have more edits than total words in the article because of vandalism and resulting article reversions.

\(^9\) I did not analyze their accuracy, but since they are largely stable and, judging from their discussion pages, consensus has been reached on their texts, it appears safe to assume that they are relatively accurate.
This sort of nonlinear development is characteristic in many instances of textual curation, particularly in distributed environments where actors must negotiate varying understandings of community norms and accepted practices. Here, the content has been filtered for relevance, for plagiarized material, and for duplication. As the articles were built out through small contributions, contributors filtered their own prior knowledge and resources for content and recomposing this unoriginal information into encyclopedic text, working in a genre and community that requires that all data be previously vetted and published so that it can be accurately cited. Wikipedia employs a “no original research” policy, requiring all material to be citable through verified, published resources (“Wikipedia:No original research”). Then, through publication, editors offer those contributions to the community for further filtration and recomposition.

**Recomposition: Fact Tagging and Content Editing**

One of the ways that the work of recomposition occurs is through editing and what is known as fact tagging. Examples include textual improvement by tagging entries that require verification and/or need additional citation in order to meet project standards. The Fact tag, which inserts a [citation needed] hyperlink after sentences that need citation, is one such tag. Doing so is not necessarily an admonition; it can also be a collaborative gesture. In late March 2009, an editor inserted this tag liberally in sections of the Minerals article, noting that this was meant as a helpful gesture: “large-scale fact tagging, typo-fixing, etc.” Six minutes later, s/he added, “another fact tag - will remove the
material if not fixed in the next day or so, but don’t want to disrupt authors ‘work in progress.’” In doing so, this editor was demonstrating one aspect the dialogic nature of curatorial work, particularly within Wikipedia: constantly working with collaborative writing and thus frequently negotiating changes rather than just autonomously making them. Instead of making the assumption that a text had been poorly written, Awickert generously assumed it was a work in progress that would continue to improve, and left a note saying that if it didn’t improve in a certain amount of time, it would be deleted in order to maintain the integrity of the encyclopedic project.

**Findability and Wikification**

Many aspects of wikification, or the process of formatting text according to community conventions, involve the strategic insertion of hyperlinks that facilitate navigation in a variety of ways. These links function as cross indexing, as taxonomic navigation, as citation, and as a path to community portals, as we saw above. Each of these functions demonstrates a distinct form of curatorial skill and engagement. Often, these links also function as keywords for the purpose of driving search engine page rank, which is also influenced by both the number of links on a page and the number of links to a page. For example, at this writing the Fortification article has been edited 48 times to add links to other article pages, and only one link was removed. It is a typically densely interwoven hypertext with multiple links in nearly every paragraph, offering the reader a quick leap to topics as diverse as Sumer, Oppida, the Yongle Emperor, and Concrete. Additionally, it offers an extensive list of other relevant pages in the “See also” section, as is typical in many established pages, and cites external sources by linking to them.

---

10 For more on composing for search engines as audiences, see Havalais, 2008.
Links to other pages within *Wikipedia* serve as digital forms of the cross-indexing seen in Western print encyclopedias since the early eighteenth century. The addition of category links further reinforces the interconnected nature of the project, but this time from an outward-in perspective. On the Fortification page, heavy editors of the topic did not insert category links. Instead, these links have been added by Wikipedians who are working on the Categorization WikiProject, a formal initiative with the goal of categorizing all articles in *Wikipedia* (“Wikipedia:WikiProject_Categories,” 2013). In the process of working through the community task list, they often by necessity deal with article topics that they’re unfamiliar with, and occasionally this leads to conflict. For example, the Falconry article is currently categorized only as “falconry,” but at one time it was also categorized under “blood sport.” The page editors, who included several experienced falconers, took issue with this categorization and removed it, leaving the note “Blood sport? wtf?” This sort of conflict is the common result of two editors who each exert their own individual agency and curatorial philosophies within a communal, participatory project. In this instance, there is little direct conflict other than a short exclamation and the removal of the offending category.

The same was also true of this exchange in the Gardens article, which features a list of gardens mentioned in famous literary works, including religious texts. The title of this list underwent some revision early on, beginning its life as “Gardens in fiction.” On April 2, 2004 the editor Fennec changed it to “Gardens in fiction, religion and myth” because of concerns about alienating audience members. S/he noted, “Watch what you call ‘fiction’, [sic] people are liable to get
upset :).” While the edit imposes change on a previous writer’s work and the tone of the note is rather terse, the note also explicitly works to avoid conflict through the use of a concluding smiley. Additionally, the edit moves the title toward a more typically encyclopedic description. There was no overt conflict in the history or discussions over this edit, but at some point in the life of the article the list title changed again. It currently stands as “Gardens in literature.” Such taxonomic issues are a common facet of curatorial work, particularly as composers work toward a navigable, stable architecture.

Another prominent category of interwiki links are links to the same article in Wikipedia translation projects. This international collaboration builds a digital network between the various language versions of Wikipedia, increasing the chances of non-English speakers locating the article in their own language after being directed to the English version by, say, a U.S. Google search.

Finally, external links in the Reference section of each article provide pathways to cited resources. These links require several sorts of curatorial labor: verifiability assessments ensure that citations are in place and linked; link rot patrols flag broken links for replacement; and both humans and bots monitor links for spam or link farming. This work, along with the link maintenance previously described, tightens the connections of a monumental network of connections to a plethora of texts and entities both within the project itself and to outside web pages. Taken as a whole, these connections coalesce into an articulation of Wikipedia’s place as a public resource that exists firmly within the open cultures of humanity rather than in any specialized educational environment.
Pedagogical Implications

Certainly, all of our students will not become writers of reference texts or crowdsourced projects like Wikipedia. But they will become writers in an increasingly networked world, and the writing they produce will be (and already is) almost entirely digital, as the Stanford Study on Writing (Lunsford, Fishman & Liew, 2013) and MSU’s “Writing Lives of College Students” study (Grabill & Pigg, 2010) have both shown. They will also inevitably deal with highly collaborative, distributed environments. Developing critical attention to curatorial aspects that facilitate rhetorical performance in both collaborative and audience-centered ways is essential in the advanced composition classroom. Designing pedagogical opportunities for students to actively build and shape complex information ecologies that go beyond individual or even small-group websites provides an important developmental experience.

This work encourages students to think actively and rhetorically about the influence that multiple aspects of their writing may have on an audience, including not just their primary texts but the paths to, through, and between those texts. Nicotra suggests as much in arguing for the value of incorporating folksonomic work in the composition classroom: “Since the implicit goal for many rhetoric-based writing courses involves training students to think rhetorically about the effects that their writing might have on an audience, incorporating folksonomy and other nontraditional composition practices into the material of the course increases the chance that students will leave the composition classroom with the unambiguous conviction that writing can be both
democratic and participatory” (274). Developing a whole-class project that offers students a chance to engage in public writing that offers value to the local community is one important aspect of teaching curation, since it encourages students to analyze concrete audiences with an eye toward civic discourse that is manifested through building ways for those users to access information.

Creating this sort of opportunity requires nothing more than willingness to carefully design a class around this sort of large and long assignment, access to free digital production platforms such as WordPress, a relevant reference project that appeals to the local student population, and a willingness to explore large-group collaboration. In order for a class to take up a project of sufficient size to really grapple with the experience of creating a curated text, the instructor may choose to shape the entire semester’s work to the work of building out such a site, which includes initial rhetorical analysis, developing the basic site architecture and helping students self-select to topics based on interest, conducting careful research, filtering, drafting, and recomposition, inserting and maintaining navigation, usability testing, and iterative review.

Making a project like this work for both the students and the instructor requires a willingness to work through not only the pragmatic processes of curatorial work but also the vagaries of collaboration. Scheduling readings about collaborative practices and time during class meetings to discuss relevant strategies and issues is vital. Students are often not used to negotiating a typical large team meeting about site architecture, for instance, but will need to be able to

---

11 For more on teaching collaborative writing in both large and small groups and a variety of digital environments, see Kennedy & Howard, 2013.
negotiate decision-making meetings in the future whether they choose to work in corporate, nonprofit, or academic organizations. Transparent coaching and discussions are key to helping them learn these skills.

Some early planning concerning the scope of the project students are asked to undertake is also vital. Will it be a stand-alone project that will be finished at the end of the term, as when a class in the UMN Department of Writing Studies worked to build a local archive of the 35W bridge collapse, or will it be an ongoing project that various classes take up over time? Will it be constructed in a single course, or will multiple classes (perhaps at multiple campuses) contribute? Since there are diverse skills to teach and collaboratively negotiate in this kind of course project, instructors may find it critical to begin with the simplest formulation: a single-class project completed in one term. Still, there is much to recommend a project that continues to see development over time, either by a repeated class or a number of classes that focus on entirely separate topics. For instance, an online guide to the local area may be maintained by a central Digital Writing class but accept contributions from classes focused on local histories and food politics. Examples of longitudinal, multi-class projects The Rhetoric and Composition Wikibook, developed by Matt Barton and his students at St. Cloud University (http://en.wikibooks.org/wiki/Rhetoric_and_Composition), the Our Michigan Avenue project at Michigan State University, directed by Jon Monberg (http://www.ourmichiganave.org/), and the [redacted] at [redacted], which I direct [URL redacted]. Many other successful formulations are, of course, possible.
Conclusion

Textual curation of complex information ecologies is no longer the future; it is simply what writing looks like now, particularly in large, open, digital reference texts. In order to prepare students for successful careers, we must continue to move toward including what Lunsford called “all the resources of a full range of media” in digital writing classrooms by finding ways to look beyond standard single-authored or small-group unit assignments that focus on essayistic digital products. As she memorably said in her Computers and Writing keynote talk,

In this scene of secondary orality and secondary literacy, student writers must be able to think critically and carefully about how to deliver the knowledge they produce. … It is as though our old reliable rhetorical triangle of writer, reader, and message is transforming itself before our eyes, moving from three discrete angles to a shimmering, humming, dynamic set of performative relationships” (2006, 170).

These performative relationships are centered in the curated product itself, in the functions that aspects of complex informative texts should perform and help users perform. The small and often invisible texts that make the world, and websites, run are essential elements of this kind of writing that help us realize the fullest potential of digital environments. These varied aspects of functional digital
composition and compositional labor help us better account for the full range of skill and labor in new media writing in our classrooms, in our research, and at our own desks. Textual curation is no longer the sole provenance of digital humanities scholars, information science experts, or digital rhetoricians. It is an essential skill set for writing teachers, writing students, and anyone who plans to work in a world that runs on digital texts.
Works Cited


Ball, Cheryl E. (2013). Pirates of metadata: The true adventures of how one journal editor and fifteen undergraduate publishing majors survived a harrowing metadata-mining project. In Stephanie Davis-Kahl & Merinda Kaye Hensley (Eds.), *Common ground at the nexus of information literacy and scholarly communication* (pp. 93-111). Chicago, IL: Association of College & Research Libraries.


Freedman, Aviva, and Graham Smart. (1997). Navigating the current of economic policy: Written genres and the distribution of cognitive work at a financial institution.” Mind, Culture, and Activity, 4.4, 238–255.


Luce, Avi. (2014). ‘It wasn’t intended to be an instruction manual’: Revisiting ethics of ‘objective’ technical communication in gaming manuals. In DeWinter, J. & Moeller, R.M.


