The Shifting Genres of Scholarly Multimedia: Webtexts As Innovation

Cheryl E. Ball
West Virginia University

ABSTRACT

In this article, I discuss webtexts as a multi-genre, multi-mediated type of research publication. After briefly describing what webtexts are, how they function, what their histories have looked like, and what media have been used to design them, I examine more closely what genres have been assigned to them and how those genres have fluctuated across time. This article explores how the genres of webtexts can be difficult to stabilize due to their technological and media innovations. This lack of stability is an unsolved preservation problem that rich metadata can partially ameliorate so that readers can interpret and interact with webtexts into the future. Although there doesn’t exist any official controlled vocabulary for metadata that would apply to webtext genres, this article outlines one attempt at creating a vocabulary for webtexts and their media elements that attends to their historical and generic shifts.

Keywords
Writing studies; Rhetoric; Digital publishing; Webtexts; Genre; Metadata; Web history

INTRODUCTION

In the last two decades, research on academic publishing has offered studies such as citation rates in and between disciplines (e.g., Bott & Hargens, 1991; Hyland, 1999; Smart & Bayer, 1986), peer-review processes and biases (e.g., Fitzpatrick, 2011; Harnad, 1996; Lee, Sugimoto, & Zhang, 2013; Smith, 2006), and the economics of publishing in closed and open-access venues (e.g., Kahin & Varian, 2000; Tenopir & King, 2000; Willinsky, 2006), among many other topics. But for the wide swath of scholars whose research questions are not focused on the production value of academic publications, much of this bibliometric literature will not seem relevant. Indeed, most scholars are likely focused on writing up the single grant application, article, or book necessary to conduct their research and
Ball, The Shifting Genres of Scholarly Multimedia

The research (in practice and in theory) is often limited to small disciplinary communities in which multimedia is already accepted as a valid and valuable method of scholarly communication (e.g., media studies), and smaller still within communities that value experimentation in their scholarly production.

Ironically, the field where readers can find most multimedia-rich scholarly publications is digital writing studies, which has a 20-year history of such work. Only more recently has media studies and art, fields that certainly value media as a rich communicative tool, made the move to incorporate media as a valued scholarly production tool. For instance, as recently as 2014, multimodal scholars Bill Cope and Mary Kalantzis wrote:

Whole disciplines limited in their publication opportunities by traditional textual exegesis – such as the arts, media and design – might formally be brought into academic knowledge systems in the actual modalities of their practice. New units of knowledge might be created at levels of granularity other than the singular article of today’s journals system; fragments of evidence and ideas contributed by an author within an article, and curated collections and mash-ups above the level of an article, with sources duly credited by virtue of electronically tagged tracings of textual and data provenance. (n.p.)

Why are so many disciplines that value multimedia hesitant to publish scholarly versions of their research that use multimedia as its main form of argumentation? I have discussed elsewhere several infrastructural reasons why disciplines do not adapt their publishing mechanisms to suit multimedia-based scholarship (see Ball, 2015; Eyman & Ball, 2015), but on a very practical level, I believe—as is the same with teaching students new genres—that scholars won’t know how to author scholarship that primarily makes meaning out of multimedia without first understanding why and how these texts work. This is a question of genre that I will discuss in the next section.

THE GENRES AND CONVENTIONS OF WEBTEXTS

Scholarly multimedia, or what I call webtexts, are research-based artifacts that enact an author’s scholarly argument through a web-based design. Each webtext is unique since each depends on the technology and media an author might need to ful-
fill her argument. Figures 1 and 2 are screenshots taken from one webtext in the online journal *Kairos* where the author designed a *wunderkammern*, or curiosity cabinet, to argue that visual arrangement facilitates the invention process in writing. Readers must interact with this wunderkammern, clicking on the cabinet to open it (Fig. 1), discovering the 36 nodes inside of it (each of which contains a selection of images, animations, and written text that attend to a smaller point within the argument; Fig. 2), and meandering through the nodes in any order they want to gain a full understanding of the webtext. (These screenshots miss the animation that is a crucial part of the argument.)

Such a webtext makes an argument that is different from the way it would be made in a traditional scholarly article. Previous scholarship that compares webtexts to traditional scholarly articles has suggested that webtexts use many of the same rhetorical moves as print-based articles do. For instance, Allison Brovey Warner (2007) remarked, “many of the webtexts published in [Kairos] address traditional scholarly goals in key conventional ways” (/tool/cat-a/convergence.html). Those conventions include content, arrangement, documentation, and tone. However, webtexts also include conventions that distinguish themselves from print-based texts, such as explicit attention to and inclusion of structural design, form/content relationship, navigational design, link strategies, node strategies, visual design, and multimedia (tool/cat-b/index.html). Warner pointed out that web-based conventions meet familiar goals of scholarship in presently unfamiliar ways. If they do meet such goals in different or innovative ways, it is important to consider what value/s they add toward creating a new genre of scholarly online text. (tool/cat-b/divergence.html, emphasis added)

In a webtext that draws on Warner’s methodology, James Purdy and Joyce Walker (2012) provided detailed examples of how webtexts meet familiar
goals of scholarship in both familiar and unfamiliar ways through

- Explicit argument (e.g., literature reviews, peer review, images as evidence, etc.)
- Speculation (in language and design)
- Implicit association (e.g., hyperlinks, juxtaposing textual elements)
- Dialogic exchange (e.g., through implicit and explicit, technological means)
- Formal enactment (through form/content relationships).

Their point was to trace the rhetorical moves that webtexts make, to show that they are scholarly, occasionally in ways—and through various media types—that differ from print-based scholarship. Indeed, they concluded that “Identifying the rhetorical moves characteristic of digital scholarship can help us begin to understand these new academic genres and their value to meaning-making in the discipline” (n.p.). Webtexts are a genre of scholarly multimedia that can be traced back to 1996, when the journal Kairos: Rhetoric, Technology, and Pedagogy first began publishing (see http://kairos.technorhetoric.net). As an editor of this journal since 2001, I have worked first-hand with over 350 webtexts through the submission, peer-review, and copy-editing processes, gaining intimate knowledge of their genres, and generic fluctuations, across time. There were other humanities-based electronic journals that existed prior to Kairos’s inaugural publication on January 1, 1996 – such as Postmodern Culture (circa 1990) and RhetNet (circa 1995) – but Kairos is considered to be the first electronic journal that has continuously and exclusively published webtext content. In addition to Kairos, other scholarly journals that exclusively publish webtexts include Computers and Composition Online (1996–2000; 2004–present); Vectors: Journal of Culture and Technology in a Dynamic Vernacular (2005–2007; 2012–2013; on hiatus as of this writing); and Journal for Artistic Research (2011–present). Each of these journals publish webtexts that follow the five criteria listed below.

- Authors design their own webtexts, either individually or in collaboration with co-authors or designers.
- The design of the webtext must convey the author’s scholarly argument and not be gratuitous.
- A webtext’s purpose is to convey new knowledge in a discipline through the presentation of research. It is not meant to function as a primarily aesthetic method of conveying creative practice, such as with examples of electronic literature or interactive multimedia for artistic effect.
- Websites can incorporate any number of design elements and media appropriate to the World Wide Web (e.g., java script, links, graphics, and, more recently, streaming media, etc.), as long as that inclusion is rhetorically in line with the scholar’s argument.
- A webtext cannot be printed without losing its argument.

This set of criteria—genre conventions of webtexts—is gleaned from my 15 years of working with Kairos, publishing webtexts in these and other journals, reading webtexts published in other
Ball, The Shifting Genres of Scholarly Multimedia

The Web is a truly multimedia medium insofar as any other medium can be embedded within a web page and linked into a Web site. As a medium, it can accommodate, in principle, and increasingly in practice, any modality and it can at least simulate most other media. It is also a hypertextual or hypermedia medium because elements in other media and modalities can be linked together in ways that allow the user to choose a variety of paths through the Web site in the course of time.

Videos, on the other hand, are exclusively linear in their presentations due to their timeline-based journnals, and teaching authors and students to compose webtexts. While each journal may have its own publishing requirements—such as Kairos’s submission guidelines that require that all webtexts be preservable on the journal’s servers and presented through an HTML infrastructure—those requirements are almost always technical in nature, not generic. Technical conventions shift with time, so I have limited the genre conventions to those that have remained somewhat rhetorically static over the last 20 years, such as the necessity of the form:content relationship as indicated in both Warner’s (2008) and Purdy and Walker’s (2012) webtexts, as well as in other scholarship that describes evaluative criteria for peer-reviewed scholarly multimedia (see, e.g., Ball, 2013; Kuhn, 2008; Kuhn, Johnson, & Lopez, 2011).

In addition to the few webtext journals mentioned above, there are a number of academic journals that publish scholarly multimedia similar to webtexts. For instance, Audiovisual Thinking: The Journal of Academic Videos (2010–2014; currently on hiatus), The Journal of Undergraduate Multimedia Projects (TheJUMP; 2010–2014; on hiatus), and the Journal of Visualized Experiments (JOVE; 2006–present) have all published videos as the primary medium of research dissemination. Video, however, isn’t a genre, it’s a medium—consider the hundreds of genres and sub-genres that reside on the video-streaming platforms YouTube and Vimeo: from far-flung examples such as cat videos (e.g., sub-genres: singing cats, cats who are jerks, cats sitting in containers) to instructional videos (e.g., sub-genres: home-improvement, classroom-based, game-play, etc.) and almost anything you can imagine in between. The ubiquity of video in the 21st century may make it seem as if video is a genre itself, but each of the genres and sub-genres mentioned above have their own set of conventions and rhetorical situations, which make them distinct—even as they are all distributed within the video medium. YouTube has a lengthy discussion of tagging videos-as-media versus tagging video subgenres. This distinction becomes important when discussing how readers interpret and interact with these texts. In the case of videos published as scholarship in journals such as AudioVisual Thinking, TheJUMP, or JOVE—each of which has its own set of genre conventions that make the videos within each radically different from the others—it is equally important to recognize the media that underlies the genres of scholarship so we can best determine how such texts should be preserved for the scholarly record.

In describing the role of media within social semiotic systems, Jay Lemke (2006) wrote, “The medium, as such, is the material technology through which the signs of the system are realized or instantiated. ... In many cases, we have only a single name for both the modality and the medium, as for example with video” (pp. 5–6). Webtexts, however, are neither a single medium nor a material technology for delivery. They tend to be infrastructurally complex, combining many media and technologies across multiple webpages or layers that are meant to be navigated hypertextually via nonlinear reading paths, as Lemke (2006:6) observes:
technical requirements; but, webtexts can contain videos as well as other media in combination for an overall non/linear reading path or trajectory, and, as Lemke said, “a trajectory across links within a Web site may already carry us across different genres and different media using different modalities” (p. 6). Drawing on Lemke’s discussion of the complexities of multimedia content, I argue that webtexts are different genres delivered using the affordances of the Web and, indeed, may contain multiple genres in a single webtext (including videos). Although Lemke (2006) was referring to surfing across websites in his discussion of genres on the Web, the same can be applied to the hypermediated nature within webtexts themselves when he said,

We are learning to make meaning across these traversals (Lemke 2002a, 2002b, 2003), in ways that are trans-generic and trans-institutional. Such traversals are still relatively free of the constraints of conventional genres and represent an important potential source of radically new kinds of meanings. (p. 7)

The exercise of separating genres of webtexts from stand-alone media types such as academic videos is not meant to make webtexts into unicorns, but to give us pause for closer analysis of these long-existing, media-rich genres. I facetiously use the metaphor of a unicorn here because of scholarly publications that have long overlooked or denigrated the existence of webtexts (e.g., English, 2005; Cope & Kalantzis, 2014). Yet, webtext practices are fundamental to some disciplinary fields. Most of the research on e-scholarship (e.g., Borgman, 2007; 2012; Cope & Phillips, 2014) focuses on STEM fields, often because it is wrongly assumed that humanistic disciplines publish books instead of journal articles. If there is any discussion of multimedia in scholarly publishing, it is often in reference to what is colloquially called an article-plus, a PDF or otherwise-linear article that includes an illustrative video or interactive dataset, literally as an aside or sidebar in some journals, such as Elsevier’s collection of Articles of the Future (http://www.articleofthefuture.com/). Commercial scientific publishers almost always present these types of articles-with-multimedia as experiments—rightly so given the unicorn status of webtexts and the general inconsistency of most webtext publishers to maintain their scholarly archive (see Ball, 2015; Eyman & Ball, 2016).

Maintaining the scholarly archive is, perhaps, the single-biggest challenge for webtext publishers since webtexts cannot be archived in databases and indices such as JSTOR and EBSCO, which only take linear texts and PDFs. Unfortunately, this means webtexts may go missing over the life of a journal. The impact of lost scholarship trickles down to future authors who will have to decide, based on the reputation of the journal, whether to submit or not. Given my focus on webtexts in humanistic contexts—in which webtext venues are primarily independent from scholarly organizations and commercial publishers—maintaining the scholarly archive within a journal publication can be burdensome. (This is a primary reason for building the academic publishing platform, Vega, which Andrew Morrison and I are working on—so that the challenge of preservation isn’t burdensome to authors, editors, or publishers.) Elsewhere, I’ve discussed best practices for sustainability in media-rich, electronic, scholarly publishing (Ball, 2015; Eyman & Ball, 2016), but those suggestions exist at the level of the journal and publisher, not at the level of the webtext or the media element, both of which concern the author–editor–reader relationship that is built through the uptake of a text. It is these levels – webtext and media element – that I will focus on next. Specifically, how does an author or editor make a webtext readable to researchers far into the future? How can such texts be preserved despite the ever-changing digital and material technologies
they employ? And, finally, how does genre impact the preservation of webtexts?

**METADATA AS AN INNOVATION MARKER**

This lack of stability for webtexts is a mostly unsolved preservation problem that rich metadata can partially ameliorate so that readers can interpret and interact with webtexts into the future. Although there doesn’t exist any official controlled vocabulary for metadata that would apply to webtexts specifically, I outline one attempt at creating a vocabulary for webtext genres that attends to their historical shifts over time.

Metadata, at a very basic level, is data about data—the finding aids of objects in an archive (Baca, 2008). Authors write metadata on their articles already—every time they craft an article and include their name, affiliation, email address, abstract, keywords, and biography, they are writing a form of metadata. That metadata helps make their article findable, indexable, and searchable. All of these functions are easy to process when the article is primarily made of words. Webtexts, however, are not primarily words (or, at least, not always): instead, they comprise a series of linked files that contain a combination of words (on the HTML pages) and other media elements (embedded in or linked from the HTML pages), such as the following breakdown:

- 8 HTML files (grouped in thematic folders, as needed)
- 10 image files in an /image/ folder
- 2 CSS files
- 3 PDFs
- 2 video files
- 1 audio file.

In generating metadata for this sample webtext, we have two levels of consideration, as mentioned above: the webtext-level and the media-element level. At the webtext level, metadata could be generated by equating such a composition to an article, which comes with some standard metadata descriptors using, for instance, a major metadata schema like Dublin Core, which has been the premier metadata standard for library and information science workers for decades. The “Guidelines for Encoding Bibliographic Citation Information in Dublin Core Metadata” (i.e., the guidelines for creating metadata for journal articles), poses the following problem for journal articles:

Many [metadata] implementations have needed a way to describe journal articles and, because of the lack of any guidelines, have invented many ad hoc solutions, which solve the problem locally but discourage interoperability. These guidelines provide a recommended and thus interoperable way to encode bibliographic citation information within a Dublin Core description. (n.p.)

Their recommendation for a Dublin-Core description of a journal article looks like Table 1 below.

There are two related notes about this example: It refers to a print-based journal article (as acknowledged elsewhere on the webpage this table was pulled from), which is, in part, indicated by the “dc:type” property (listed as Text in this table). DC:type refers to one of the 15 properties used in the Dublin Core’s Metadata Element Set. Type, in that core set’s list of elements, is defined as “the nature or genre of the resource” (n.p.). It is further recommended that type be used with “a controlled vocabulary such as the DCMI Type Vocabulary [DCMI_TYPE]” (n.p.), which “provides a general, cross-domain list of approved terms that may be used as values for the Resource Type element [of Dublin Core] to identify the genre of a resource” (see DCMI Type Vocabulary, 2006). Unfortunately
Table 1. Dublin Core description of a printed journal article.

<table>
<thead>
<tr>
<th>Property</th>
<th>Encoding Scheme URI</th>
<th>Value String</th>
<th>Value URI</th>
</tr>
</thead>
<tbody>
<tr>
<td>dcterms:creator</td>
<td><a href="http://purl.org/dc/elements/ns/creator">http://purl.org/dc/elements/ns/creator</a></td>
<td>2671</td>
<td></td>
</tr>
</tbody>
</table>

that list of genres is purposefully broad so that it can serve the widest swath of objects possible. The list includes the following genres:

- Collection
- Dataset
- Event
- Image
- InteractiveResource
- MovingImage
- PhysicalObject
- Service
- Software
- Sound
- StillImage
- Text

For example, an Image DCMI Type can include “images and photographs of physical objects, paintings, prints, drawings, other images and graphics, animations and moving pictures, film, diagrams, maps, musical notation. Note that Image may include both electronic and physical representations” and yet there is also a MovingImage [sic] type that includes “animations, movies, television programs, videos, zoetropes, or visual output from a simulation. Instances of the type Moving Image [sic] must also be describable as instances of the broader type Image” (n.p.). The Dublin Core Metadata Initiative acknowledged in several locations on its massive schema website that “the terms in DCMI vocabularies are intended to be used in combination with terms from other, compatible vocabularies in the context of application profiles” (Dublin Core Metadata Element Set, n.p.).

Webtexts, as genres that contain many subgenres, would be categorized as DCMI Type InteractiveResource—“A resource requiring interaction from the user to be understood, executed, or experienced” such as “Web pages, applets, multimedia learning objects, chat services, or virtual reality environments” (DCMI Type Vocabulary, n.p.)—not as Text, like the print journal example mentioned above. Within either InteractiveResource or Text, as with Image, there can be multiple subgenres to categorize. This further classification can be especially important when it comes to webtexts,

1 Dublin Core does offer an example of metadata usage for an electronic journal article, but that example mysteriously excludes many of the print-article properties that would be needed for an e-journal article, too, including Publisher and Type.
since the subgenres these scholarly texts embrace might include such wide variety as video manifesto (Miles, 2015; Watkins, 2008), parody (Theamishagur, 2009; Walls, 2008), game (Nelson et al, 2013; deWinter, 2014), or comic (Jacobs, 2014; Johnson, 2014; Parish, 2012). Webtexts require more description of genres than metadata standards such as Dublin Core offer, especially given the constantly changing affordances within digital technology as seen through changing Web standards, updated media types, and obsolete software. When all of these changes come together in a single webtext—such as the use of Macromedia Director (as an early example) or Adobe Flash (as a recent example), which were once-stable design platforms for journals like Kairos—both authors and editors need to consider crafting more robust metadata than the descriptors available in InteractiveResource. Dublin Core and other descriptive forms of metadata (including the Library of Congress’s MARC’s Genre Term List, Moving Image Genre Form Guide, and Thesaurus for Graphic Materials) are really only meant to help readers locate objects, and perhaps to know what hardware or software emulator to use on a MIME Type that is no longer playable in most browsers (e.g., a Shockwave executable file or RealPlayer file). As if to squelch the concern of multimedia publishers everywhere, Murtha Baca (2008), editor of the foundational textbook on metadata, reminded us:

There is no “one-size-fits-all” metadata schema or controlled vocabulary or data content (cataloging) standard. Institutions must carefully choose the appropriate suite of metadata schemas and controlled vocabularies (including collection-specific thesauri and local picklists), along with the most appropriate cataloging standards (including local cataloging guidelines based on published standards) to best describe and provide access to their collections and other resources. (principles.html)

3 A MIME identifies grouping of filetypes such as text/css, text/html, application/pdf, image/gif, image/png. Mimetypes are a controlled vocabulary created by IANA – a full listing can be found here: http://www.iana.org/assignments/media-types/media-types.xhtml. Mimetypes are not directly indicative of genre, but they lend help in pinpointing the mixtures of genres that might be embedded in a webtext.

Inclusion of metadata should be an integral part of the invention and production process for digital works on the part of authors as well as a standard element of the publishing process. “The active construction of metadata should be a compositional practice because it is emblematic of the ways an author deploys rhetoric, design, and code as the means by which a given webtext engages scholarly, social, and technical infrastructure” (Eyman & Ball, 2014, p. 116). In the next section, I describe one effort to create a robust genre description for Kairos webtexts that would be usable for finding and preserving these innovative digital objects.

INSTABILITY AND INNOVATION IN WEBTEXTS

Starting in Fall 2010, as part of a National Endowment for the Humanities Digital Humanities Startup grant, I led a group of faculty and student researchers through a metadata-mining project for Kairos webtexts. Working from a Dublin Core (DC) and Open Journal Systems (OJS) inspired metadata schema created by then-Kairos Associate Editor Kathie Gossett, Senior Editor Doug Eyman, and myself, 19 researchers captured 32 fields of metadata at the webtext and media-element level. Min-
Some of those webtext- and media-specific properties are listed here:

**Webtext-Level:** Author, Title, Abstract, Volume & Issue, Section, Date Published, Peer-Reviewers, Peer-Reviewed Status, Publisher, Language, DOI, DCMI Type, URI, Genre, File Size, Creator

**Media-Element-Level:** FileName, MimeType, Page Title, Alt Text, DCMI Type, URI, Genre, File Size, Creator

**Authors:** Affiliation, Rank, Email.

Many of these properties correspond to the Dublin Core terms and others from DC and OJS that we added, although in most cases, the names of the terms were changed for convenience since I was working with 15 undergraduates unfamiliar with **Kairos**, webtexts, or metadata schema. For instance, the term Author mapped onto DC’s Creator term, Abstract to DC’s Description, Section to Source, and so on. Some properties, like MimeType and DCMI Type were so specific as to warrant not changing for the student-researchers, particularly since there were controlled vocabularies involved. While I have described elsewhere how this project was undertaken (as a pedagogical tool; see Ball, 2012), here I focus on discussing the Genre properties at the Webtext level and the Media-element level, which proved to be the single-most difficult property to determine in this project.

At the Webtext level, Genre was slightly easier to map, as it often corresponded to the Section of the journal in which the webtext appeared, although combing through the archives of the journal made this process more difficult. The scope of the mining project included retrieving or creating metadata for all back issues of the journal; at the time we started the project, we were at Volume 15, Issue 2, which translated to 15 years of publication with at least two, if not three, issues per year. The names of journal sections had changed over the years. For instance, sections variously called **Kairos Interactive**, Responses, and Disputatio all published webtexts that could be labeled as a “response” genre. Similarly, the sections called Classroom Spotlight, Pixelated Rhetorics, Praxis, and PraxisWiki could have discovered all sorts of problems with the way webtexts, and elements within webtexts were and, for the most part, were not, cited. He was attempting to create a remix engine so that readers could better use the existing media assets the journal had already published, which we hoped would also ameliorate the citation problem by providing DOIs (or at least directly accessible file names) to media elements that users might employ.

4 By that point, Senior Editor and digital rhetoric scholar Douglas Eyman had been researching citation practices of webtexts for nearly a decade (see 2000, 2007a, 2007b), and
comprise the genre “praxis,” which would be webtexts focused on theorized classroom practice. Other sections, such as Interviews, News, and Reviews were easier to categorize as genres according to their section titles, and these sections rarely (if ever) changed names. We did create and assign subgenres to the peer-reviewed sections of the journal, and those were, concomitantly, the sections that were less stable in terms of Section names. For example, the “feature” genre could be found in the Features, Topoi, or CoverWeb sections, for which sub-genres such as case study, course study, teaching narrative, editorial, and others might be listed. We created a controlled vocabulary for all sections by discussing the rhetorical aims of several example webtexts in a norming session and proposing generic labels for those aims. Below is a partial list of webtext subgenres (with brief descriptions):

- editorial – editor’s statements
- discussion – town hall discussions
- book review – review of a book
- software review – review of software/CD, etc.
- review essay – review of multiple objects (includes micro-reviews)
- interviews
- chat interview – interview conducted in a MOO, IRC, chat, etc.
- survey
- case study
- course study
- assignment study
- reflection
- conference review
- listserv review
- teaching narrative – (includes narratives in the PraxisWiki section)
- ethnography
- usability study (of space, of website design...).

These are not further broken down by Section or main genre since some of these subgenres could cross into several sections. For instance, the discussion subgenre could appear in a CoverWeb, since Kairos used to publish special issues from the Computers & Writing conferences, at which a Town Hall was a main session. It could also appear in an old section called Kairos Meets The Authors, which were part-interviews/part-discussions between a staff member and an author who had published a webtext in the same issue. These subgenres were admittedly weird, and the student-researchers could be wildly inaccurate in applying them (so it goes with metadata...).

Our larger concern as an editorial team, however, was with assigning genre to the media elements. After all, abstracts and other webtext-level metadata such as Title, would help future researchers trace the historical importance of this journal. At the level of the media element, we wanted authors and researchers to be able to determine which videos, for instance, were instructional, performative, student/classroom projects, or something else. In this discipline, instructors often assign multimedia projects, and we are constantly looking for examples to share with students. It is an oft-repeated request of practitioners in the field to ask someone to build them a database that would help them collect student-built projects of various genres and media.

5 The reasons why sections changed their names is beyond the scope of this article, although it often had to do with a combination of editorial staff changes and shifts in the disciplinary discourse.

6 It should be said that the students wrote abstracts for all 1200+ webtexts they found. Asking for abstracts had not been a practice of Kairos until we undertook this metadata-mining project.
We also wanted to provide data on the history of MimeType usage for researchers who may eventually study the archives of Kairos to trace research questions such as how the rise of CSS changed the interface designs and sustainability of webtexts. Assigning MIME Types was easy, given that it is an Internet Assigned Numbers Authority (IANA) controlled vocabulary (e.g., a file extension of .avi will always be named Audio Visual Interleave (AVI), which will always be a MIME Type of “video/x-msvideo” just as a file extension of .htm or .html will always be named HyperText Markup Language with a MIME Type of “text/html”). However, as with the example of video from earlier in this article, a MIME Type of video/x-msvideo doesn’t tell us what the genre of that video is, even as it tells us what its file format is (and, thus, what potential codec or emulator we might need to read it in the future). Still, we want to know the genre: student project (and then what sub-genre)?, informational video?, performance?, manifesto?, and so on.

Another example we kept returning to was that of the MIME Type image/gif. It won’t be surprising to historians of the Web that gifs used to run rampant in the early days of Kairos; authors used that MIME Type for all sorts of purposes—as title graphics with a decorative font (pre-font APIs), as navigation buttons, as illustrative decorations (despite the a-rhetoricity of such usage; this was definitely a period-specific choice in Web design), as animated gifs, and more. Listing all of these genres as simply Image or MovingImage would not do justice to the rhetorical complexities of these media elements. So, as we did with the Webtext-level genre property, the student-researchers and I normed our terms for all media elements that were represented in Kairos. We sorted by DCMI Type, then MIME Type, and then looked at the file a-contextually (and then contextually, in relation to the webtext page, if needed) to determine what sub-genre it should have (see Fig. 3 for a portion of one Excel sheet we used to capture the metadata for StillImages). The controlled vocabularies we created for three media types are presented below:

### Images

- title header—title text made into graphic (may include webtext logo)
- title graphic—graphic used alongside a title (repeated as a logo)
- title animation—animated graphic as title
- splash title—graphic that readers have to click through to get to main webtext

- author photograph
- family snapshots
- illustrative photographs
- background image
- image map—an image used as the main navigation in a webtext
- book cover graphic
- navigational graphic—arrows, etc.
- clip art—graphics used for illustrative or placeholder purposes
The original version of this definition read “any still image that moves in an annoying way.” One of the reviewers for this article rightly questioned who had placed the value of “annoying” onto the media element. The provenance speaks to an interesting historical artifact within the journal: the students and I would have been talking about late-1990s animated gifs such as flying envelopes, opening mailboxes, and blinking under-construction signs, which were de rigueur for that time period in Kairos. These kinds of animated gifs are certainly seen as annoying by today’s web design standards. I have revised the definition here to make it more useful across time, since animated gifs have returned to prominence via memes and reaction-type uses.
The purpose in having a more fine-grained description of Genre applied at the webtext level was intended to help future researchers reverse-engineer the kinds of experimental scholarship the journal publishes, in cases where codecs, plug-ins, or even the Internet were no longer available. Given that the Library of Congress (LOC), who would otherwise seem to be the ultimate arbiter of such namings, had no dog in the scholarly multimedia fight at the time,8 we decided to create our own schema and be responsible for our own preservation until such time some other organization came up with a better preservation plan. The journal has been lucky to publish for 20 years (as of this writing), but I am often asked where, or what, the journal will be in another 20 years, and I do not have an answer for that. But I do know that regardless of the experimental and innovative media usage Kairos has exhibited in its history—some of which can be extremely difficult to preserve and play even just a few years after it was published—that some future researcher studying in the discipline that Kairos publishes in (e.g., digital writing and rhetoric, and related fields) will be able to trace back the field’s history in enough detail to understand that, at one point, digital writing studies valued, as one example, reviews about listservs and conferences, the latter of which we still regularly publish. Naming genres—and the difficulty of doing so when the subgenres of webtexts have changed alongside media types every few years (a topic I wish I had room to discuss here)—is also important for helping potential authors grapple with composing webtexts and helping readers, including hiring and promotion evaluators, understand the range and value of digital work being produced by scholars to date.9

8 Since 2010 the LOC has been working on more schema that would account for born-digital work (although not scholarly multimedia specifically), under the name of the National Digital Stewardship Alliance. In 2015, the NDSA transitioned to a new institutional home under the Digital Library Federation, where more work towards digital-asset preservation will continue to take place.


**CONCLUSION**

Although this project was intended simply to supply data for a database so that Kairos could migrate to a content-management system, we realized along the way what a rich collection of research questions it created. With well over a million points of data on over 1,000 webtexts (and counting), this metadata-mining project contains the possibility of so many research avenues that we made the data publicly available in the Stasis (archive) section of Kairos: http://technorhetoric.net/stasis/2016/kairos-metadata/.

The mining project also revealed significant challenges that journals face when publishing webtext genres. Most readily is the lack of long-term accessibility—both in the sense of disability issues and access and preservation issues. These two issues are intimately related, as we describe in the multi-authored webtextual white paper, “Access/ibility: Access and Usability for Digital Publishing” (Eyman et al, 2016). For Kairos’s immediate purposes in regards to the metadata mining project, we discovered the journal has a sad lack of alt attributes on images and other media-rich, non-textual file types throughout, despite submission requirements calling for alternative descriptions...
as well as editorial processes that require adding alt text when it is missing. Essentially, we require written descriptions embedded in the code on anything that is nonlinguistic text in a webtext, so that readers with impaired vision or hearing, differing neurocognitive abilities, or slow connection speeds can access metadata about these media assets in the easiest way possible. The journal was admittedly bad at mandating alt attributes and similar properties in its first decade, but by the beginning of the second decade, it had put workflows in place to ensure these descriptions were made available prior to publication.

However, as the mining project shows, this wasn’t always done as thoroughly as it should have been. Not only does a lack of accessibility metadata hurt our readability; it also hurts our findability, searchability, and sustainability. For instance, some media types, as I mentioned earlier, are no longer easily accessible through browsers (e.g., Flash and Shockwave files). Alt text would function as a metadata stand-in that could potentially describe the source (and not just the MIME and DCMI Type, but the actual genre and, thus, purpose of the file) for that missing media element, making these media elements more sustainable in terms of future research purposes—but only if that metadata exists in the first place. Figure 4, for instance, shows a screenshot of the metadata collection for an early issue of *Kairos*, in which nearly all of the alt attributes are missing. This makes all of these media files inaccessible to multiple user types and, further, makes these images impossible to find through a search. This is highly problematic in an open-access webtext where the design of the piece (including all of its media, interface, layout, etc.) should convey its scholarly argument through aesthetic and rhetorical means just the same as any written content should.

The Page Titles in Figure 4 also show how haphazard these descriptive elements were in early issues. Starting with the 2008 “New Design Debut” issue of *Kairos* (which was also the issue we began incorporating DC metadata into all HTML headers), we began assigning standardized <title> tags to every HTML page in a webtext: Kairos Vol.Issue: Authorslastname (et al), Short Webtext Title – Web Page Content Title (e.g., “Kairos 13.1: Yergeau et al, Expanding the Space of f2f – Home”). A major part of the 2008 redesign involved removing any semblance of frames from our previous interface designs, which meant that webtext-specific page titles were now visible in the browser toolbars (see Figure 5). Adding in standardized page titles also meant that readers could more easily determine what they were looking at, in terms of genre, in a given browser window (i.e. that they were reading an online scholarly journal, and a specific page of a webtext in that journal), which we felt was crucial given the lack of scholarly navigational signposting some online journals practiced. We had seen authors cite webtexts from other major journals in our field as if they were personal websites instead of peer-reviewed scholarly articles because (we guessed) there was no indication on the webtext itself or within the webtext’s metadata that it belonged to a peer-reviewed scholarly journal.

Given the innovative nature of media-rich webtexts, it behooves editors and authors of these and similar kinds of scholarly multimedia and interactive designs to include rich descriptions of their work, at the level of the work if not at the level of its parts, as I have described here. Metadata exists not only to protect media elements and article-length projects, but also the journals themselves. Many online journals have lost their collections due to bad planning because they assumed webtexts were static publications or did not need maintaining in ways that are unique to digital media-rich publications (See Ball, 2015; Eyman & Ball, 2016). As Cope and Kalantzis (2014) articulated:
Most journals are still making PDFs, still bound to the [stable] world of print-lookalike knowledge representation, but a reading of technological affordances tells us that we don’t have to replicate traditional processes of knowledge representation – digital technologies allow us to do more than that. (n.p.)

Indeed, webtexts do ‘more than that.’ But because its genres change with each new technology, it is hard for readers and researchers to notice, let alone trace, webtexts’ stability-in-innovation. By examining the multiple genres (as well as MIME Types, file sizes, etc.) through one journal’s history, we can look back at how one discipline makes and values knowledge in its fields and anticipate better ways to preserve and promote webtexts for readers and researchers in other disciplines, now and into the future.

Figure 4. Metadata screenshot showing, in the dark grey box, missing alt attribute information for a series of image files, in an early issue of Kairos.

Figure 5. Webtext page titles show up in browsers, allowing readers to more easily know the provenance of a webtext at a glance.
Unfortunately, it is outside the scope of this article to delve into how this project helped us create accessibility and sustainability guidelines for webtexts (and similar kinds of digital media and digital publishing projects). If you are interested in that project, I encourage you to read “History of a Broken Thing: The Multi-Journal Special Issue on Electronic Publication” (Eyman & Ball, 2016), which traces technologically infrastructural-wide problems with digital media publishing and how to ameliorate them, and “Access/ibility: Access and Usability for Digital Publishing” (Eyman, Ball, Boggs, Booher, et al, 2016), which offers a set of best practices for digital publishing with attention to the relationship between open access, accessibility for diverse users, and access and preservation standards. It may be enough for some readers to follow these suggested research threads, but this body of research on the innovative and difficult nature of webtexts—technologically, infrastructurally, and, as demonstrated in this article, generically—also requires new publishing practices to be put into place in order for this kind of scholarly innovation to continue. It is because of that need that Andrew Morrison and I, along with consulting researcher—editor Douglas Eyman, began work with a development studio in Oslo, Norway, in early 2015 to build Vega, an academic publishing platform that would support the generic innovations webtexts promote. Vega provides author tools, editorial workflows, and publishing features (including flexible metadata schema) that showcase webtexts as a priority genre in scholarly publishing. To follow the progress on Vega, due for release in Fall/Winter 2017, please visit http://vegapublish.com.

ACKNOWLEDGEMENTS

The National Endowment for the Humanities funded the metadata mining project with a Level II Digital Humanities Start-Up Grant in 2010. The metadata was mined by Sarah Chance, Lyndsey Eagle, Meghan Engel, David Gaudio, Susan Grogin, Melissa Hermann, Laura Patrick, Kira Plotts, Valerie Romack, Constance Ruholl, Kali Shevlin, Ameliah Tawlks, Amy Thomas, Kirsten Van De Veer, Jessica Wosniak, Gina Cooke, and Kristi McDuffie. Many thanks to Gunnar Liestøl who initially prompted me to take a genre-innovation approach to webtexts. Thanks also to the U.S.–Norway Fulbright Council, for the time and resources to work on this project, and the Oslo School of Architecture and Design for the space to complete it as a Fulbright Scholar. Finally, thanks to the Andrew W. Mellon Foundation for supporting the Vega project, and to Bengler design studio, whose work on Vega has made our research on sustainably publishing webtexts a reality.
REFERENCES


Ball, Cheryl E. (2013b). Pirates of metadata or, the true adventures of how one editor, fifteen undergraduate publishing majors, and 25,000 media elements survived a metadata mining project. In Stephanie Davis-Kahl & Merinda Hensley (Eds.), *Extend and unify: Outreach and education for scholarly communication and information literacy programs*. Chicago: Association of College and Research Libraries.


