DHQ: Digital Humanities Quarterly

2012 Volume 6 Number 1

Comic Book Markup Language: An Introduction and Rationale

John A. Walsh <jawalsh_at_indiana_dot_edu>, Indiana University

Abstract

Comics, comic books, and graphic novels are increasingly the target of seriously scholarly attention in the humanities. Moreover, comic books are exceptionally complex documents, with intricate relationships between pictorial and textual elements and a wide variety of content types within a single comic book publication. The complexity of these documents, their combination of textual and pictorial elements, and the collaborative nature of their production shares much in common with other complex documents studied by humanists — illuminated manuscripts. artists' books, illustrated poems like those of William Blake, letterpress productions like those of the Kelmscott Press, illustrated children's books, and even Web pages and other born-digital media. Comic Book Markup Language, or CBML, is a TEI-based XML vocabulary for encoding and analyzing comic books, comics, graphic novels, and related documents. This article discusses the goals and motivations for developing CBML, reviews the various content types found in comic book publications, provides an overview and examples of the key features of the CBML XML vocabulary, explores some of the problems and challenges in the encoding and digital representation of comic books, and outlines plans for future work. The structural, textual, visual, and bibliographic complexity of comic books make them an excellent subject for the general study of complex documents, especially documents combining pictorial and textual elements.

Science Exhibit

Experiments in Radio-Activity

> Open to the

Public

(Lee & Ditko. Amazing Fantasy #15, 1962.)

1

Introduction

This study provides an introduction and rationale for the development of Comic Book Markup Language, or CBML, an XML^[1] vocabulary for encoding multiform documents that are variously called comics, comic books, and "graphic novels" ^[2] as well as other documents that integrate comics content^[3] or that share formal features with comics content. A markup language is a set of machine-readable textual codes, or "tags," that are used to identify structure, semantics, and other features of documents and data.^[4] The application of these codes to a document is typically a necessary stage, often the most crucial and informative stage, of editing, analyzing, indexing publishing, visualizing, and otherwise studying or manipulating texts in digital environments. The act of encoding a document is a form of discovery, or

prospecting, in which the encoder maps a document's structure, identifies semantic elements of interest, and documents relationships internal and external to the document. Scholarly encoding is a form of both reading and writing. The reading, shaped by the constraints of a markup language, is inscribed upon and embedded within the digital text. As literary scholar and digital humanist Jerome McGann has noted, "When you mark up a text you are ipso facto reading and interpreting it. A ... text marked up in TEI [a scholarly encoding language] has been subjected to a certain kind of interpretation" [McGann 2001, 143]. Sperberg-McQueen, Huitfeldt, and Renear assert that markup is constitutive of meaning, markup is interpretive, markup is performative, markup acknowledges or licenses inferences about the text:

Markup is inserted into textual material not at random, but to convey some meaning.

An author may supply markup as part of the act of composing a text; in this case the markup expresses the author's intentions, e.g. as to the structure or appearance of the text. The author creates a section heading, for example, by creating an appropriate element in the document; the content of that element is a section heading because the author says so, and the markup is simply the method by which the author says so. The markup, that is, has *performative* significance.

In other cases, markup is supplied as part of the transcription in electronic form of pre-existing material. In such cases, markup reflects the understanding of the text held by the transcriber; we say that the markup *expresses a claim* about the text. The transcriber identifies a section heading in the pre-existing text by transcribing it and tagging it as a section heading; the content of that element is a section heading if the transcriber's interpretation is correct, but other interpreters might disagree; it is plausible to imagine discussions over whether a given way of marking up a text is correct or incorrect.^[5]

In the one case, markup is constitutive of the meaning; in the other, it is interpretive. In each case, the reader may legitimately use the markup to make inferences about the structure and properties of the text. For this reason, we say that markup *licenses certain inferences* about the text. [Sperberg-McQueen, Huitfeldt, and Renear 2000, 11]

Julia Flanders' discussion of scholarly text encoding privileges the role of the researcher/encoder and likens the encoded text to the scholarly article:

Perhaps we need to look to the pleasure of mutability. To recuperate XML, politically and aesthetically, we should be looking not to the paradigms of XML usage that arise from librarianship and from industry-level ideas of the separability of form and content, but rather to paradigms of performance of a different kind. By shifting our view we can understand XML as a way of expressing perspectival understandings of the text: not as a way of capturing what is timeless and essential, but as a way of inscribing our own changeable will on the text — in other words, as a form of reading. Seen this way, XML's presentational flexibility derives not from a separation of presentation and content, but rather from the shifting vantage points from which the text appears to us, the shifting relationships that constrain our understanding of it, the adaptability and strategic positioning of our own readerly motivations.

Ironically, this is a view which emerges most clearly at the margins of current digital text practice. It is not visible in the large digital library projects, whose workflow has come to resemble an industrial operation complete with offshore outsourcing, detailed division of labor, reliance on automation and robotics, and an emphasis, in the output, on uniformity and quantity (thankfully planned obsolescence has not yet become part of the strategy). But we can find it in the small projects designed by individual faculty, typically in conjunction with their teaching, to create digital versions of individual texts which serve as readings: often idiosyncratic, unscalable, representing private insight. They function more like an article than an archive, as a local, contingent expression of insight. [Flanders 2005, 60–61]

The development of a markup language that can support such scholarly reading and interpretation necessitates a careful study and analysis of the content, structure, and semantics of the class or classes of documents for which the language is designed — in this case, the comic book.

CBML is based on the *Text Encoding Initiative P5: Guidelines for Electronic Text Encoding and Interchange*. The *TEI Guidelines* are a mature conceptual model for digital representation of multitudinous and disparate document types: inscriptions and papyri; illuminated manuscripts; authorial holograph manuscripts; correspondence; printed books of prose, verse, and drama; critical and scholarly editions; born-digital documents; and more. The *TEI Guidelines*

make recommendations about suitable ways of representing those features of textual resources which need to be identified explicitly in order to facilitate processing by computer programs. In particular, they specify a set of markers (or tags) which may be inserted in the electronic representation of the text, in order to mark the text structure and other features of interest. [TEI 2010c]

The *TEI Guidelines* are widely used in the digital humanities and academic library communities and are maintained by the TEI Consortium, an international body modules, including modules for general categories of documents, such as prose, drama, verse, and dictionaries.^[6] The *Guidelines* also provide additional modules that address more specific textual features and metadata requirements, such as names and dates, manuscript description, linking, textual criticism, and so on. And in their most recent incarnation, the *Guidelines* provide elements and attributes for linking transcriptions to facsimile page images. This latter feature is especially useful for encoding comics and other graphics-intensive works. From these many available modules, one selects a subset that meets the needs of a particular document, project, collection, or analytical approach. The *TEI Guidelines* are extremely flexible, providing a vocabulary and mechanisms for encoding and describing a rich diversity document types. However, recognizing that not every document type and representational requirement may be anticipated, the *Guidelines* provide a well-documented system for customizing and extending the provided tag set with new and modified elements and attributes.^[7] TEI, as delivered by the TEI Consortium, is remarkably well-suited to encoding many aspects of comic books; nevertheless, conceptual clarity and practical benefits may be gained from some modest modifications and additions to the stock *TEI Guidelines*. Hence CBML, a TEI customization with elements and attributes for encoding many of the structures and features found in comic book documents.

The Book of Comic Book Markup Language

The language under discussion here is called Comic *Book* Markup Language in part to highlight the book-ness and bookishness of these documents, their material properties and bibliographic characteristics. Graphic narratives typically manifest as "books," stapled or otherwise bound leaves, perhaps thirty-six pages, with an interesting and complex structure, incorporating the graphic narrative — the sequential art and text or "comics" content — alongside a rich assortment of paratexts: advertisements, fan mail, and so on. The emphasis on both "comics" and "books" in the title of the language signals an awareness of the full range of content in the material artifact and the integration of comics content with related paratextual content. The material properties of the book — the codex form, the leaves and pages, the physical properties of the paper — are inseparable from the structure, pacing, and design of the narrative. Certainly less page-bound organizational and compositional frameworks are possible, such as the newspaper strips with long-running narrative arcs, in which the daily "strip" of three or four panels is the basic structural unit.

The traditional grouping of panels into deliberately composed groups (often corresponding to the physical page or the "strip" of a newspaper daily) is being challenged by changing publishing and reading technologies. iPhones and other smartphones have become popular devices for reading newly published comics as well as "reprints" of older comics. However, a full-page grid of panels is not easily readable on the smaller screen of the typical smartphone, so the software interfaces on such devices focus on a single panel at a time. The deliberate juxtaposition of graphic and textual elements in the original composition is shattered by the interface requirements and limitations of the reading device. As new comics work is increasingly targeted at such digital platforms, the traditional grouping of panels into compositional units resembling pages may be abandoned for new compositional strategies. Larger format devices like the iPad and

3

4

5

other tablets are better able to represent full-page compositions of panels while also allowing zooming in to focus on individual panels. The isolation of a panel from its surrounding context is not easily achieved in print media. While the reader's eyes and attention may focus on a single panel at a time, other panels on the page remain in the reader's field of vision. The migration of comics content to digital reading devices, and the structural and aesthetic implications of that migration, call attention to the impact of the material characteristics of the comic book document.

Comic books are often very formally self-conscious documents and express a fascination with their own bibliographic identities — creators (comic book writers and artists) become characters in the narrative, editorial notes refer to episodes from prior issues, comic books and paratextual elements, such as advertisements, are parodied within the comic book narrative, publication milestones (such as the first, fiftieth, or one hundredth issue are highlighted and celebrated. These literary, rhetorical, and commercial moves point to a self-awareness of the comic book as document and bibliographic object.

CBML is intended primarily for representing, modeling, and analyzing twentieth- and twenty-first-century comic books, daily comic strips, longer narratives or "graphic novels," and Web comics and other comics content published on digital platforms, such as smartphones and tablets. CBML may also serve as a possible solution for encoding certain documents we might not normally characterize as comics or comic books, but which share many formal characteristics with comics. In his influential *Understanding Comics*, Scott McCloud's definition of comics encompasses Hogarth's narrative picture series, the Bayeux Tapestry, and pre-Columbian picture writing as found in the Codex Zouche-Nuttall [McCloud 1993, 10–17].

Goals and Motivations

Comic books and graphic novels have been the subject of serious critical attention for some time, and increasingly so with the emergence of scholarly disciplines such as popular cultural studies and new areas of interest in traditional scholarly fields such as English and American literature. In 1992, Art Spiegelman won the Pulitzer Prize for his Maus, a comic book narrative of holocaust survival. In 2001, Michael Chabon won the Pulitzer Prize for his novel The Amazing Adventures of Kavalier & Clay, which relates the experience of two Jewish cousins working in the nascent comic book industry at the beginning of WWII. Comic books and the mythologies they have spawned continue to be a vital part of our popular culture and national consciousness. Witness the surprising and almost unprecedented popularity of the recent spate of superhero feature films based on the characters of Marvel and DC, the two largest publishers of comic books. In the 70s, 80s, and 90s, the Superman and Batman film franchises produced regular blockbusters. Since the more recent financial and critical success of Marvel's X-Men and Spider-Man film franchises, the film industry has been flooded with films based on comic books, many based on mainstream super hero comics, others adapted from realistic, personal, and autobiographical graphic narratives, often published by smaller publishers. In 2012, Marvel's The Avengers film from Walt Disney Studios broke all previous box office records with over \$200 million in ticket sales for the biggest opening weekend of all time [Barnes 2012].^[8] These few examples, which are frequently discussed in scholarly articles, monographs, and in school and university classrooms, demonstrate the continuing and perhaps increasing importance of comic books as an art form and cultural touchstone.

Like the study of other popular art forms — film, television, jazz, rock — comics scholarship is a thriving field of academic research. Scholarly books about comics are being published by university presses^[9] and articles about comics appear frequently in peer-reviewed journals such as *The Journal of Popular Culture*. In 2008, *English Language Notes*, published an issue devoted to "Graphia: The Graphic Novel and Literary Criticism" [Kuskin 2008]. *ImageTexT* is a peer-reviewed online journal devoted to the study of comics. In 2007 *ImageTexT* published a special issue on "William Blake and Visual Culture," recognizing the connection between modern comics and earlier art forms that integrate text and image [Whitson 2007].

In light of the ongoing cultural and scholarly relevance of comics, one goal of CBML is to support the study and analysis of comic books in the way that digital collections of, for instance, English poetry or American fiction support the study and analysis of more traditional literary forms. A large corpus of digitized comic books, along with encoded transcriptions and descriptive metadata, would allow scholars to search the text of comic books, search for keywords

10

11

7

related to topics of interest, search for the appearance of particular characters, or search for works by particular writers and artists. Additionally, when exploited to its full potential, a large CBML collection would allow searching — and other forms of computer processing and computational analysis — based on structural, aesthetic, and informational and documentary features peculiar to the genre of comic books.

12

Large digital collections of comic books would support the types of searching that is now taken for granted in large digital collections of literary and other texts. This sort of functionality has proven incredibly useful and has transformed the ways in which scholars and students conduct research and "read." This now commonplace functionality though can be particularly significant for the study of comics, given the complexity of the fictional universes that have developed in the works of many comics publishers. For instance, major characters like Superman, Batman, Spider-Man and Wolverine appear regularly in multiple titles. DC Comics' Batman is the featured character in *Batman, Detective Comics*, and other titles; he is a prominent member of the Justice League superhero teams featured in a number of different titles; and he shares adventures with other heroes in titles such as *The Brave and the Bold* and *World's Finest*. Since Batman shares the same fictional universe with most of DC's other characters, he will make frequent "guest appearances" across their entire line of publications. One will even find publishers occasionally cooperating to merge their fictional universes and publish works that feature, for instance, Superman from DC and Spider-Man from Marvel. Therefore, a scholar researching the development and representation of a particular character cannot confine herself to a small number of publications, but must consider almost the full output of a publisher or even multiple publishers.



Figure 1. From 1976, *Superman vs. the Amazing Spider-Man*, the first cross-over comic book featuring characters from comic book publishing giants DC (Superman) and Marvel (Spider-Man). [Conway 1976]

research collections that in turn support different types of searching and digitally-enabled analysis. Another motivation behind the development of CBML is the desire to explore more generally the modeling and representation of the broader class of documents that tightly integrate pictorial images and text. Comic books are just one such type of complex graphic document; other examples include illuminated manuscripts; seventeenth-century alchemical manuscripts, with hand-drawn figures and graphic symbols; artists' books; artists' sketchbooks; illustrated poems like those of William Blake; letterpress productions like those of the Kelmscott Press; illustrated children's books; newspaper and magazine advertisements; and even Web pages and other born-digital media. Comic books provide some of the more obvious examples of this integration and co-dependence of image and text, but lessons learned in the modeling of comic books may be applied to other forms.

Martha Rust, in her essay " 'It's a Magical World': The Page in Comics and Medieval Manuscripts," writes:

Both medieval book artists and contemporary cartoonists make use of the page as a device for giving their readers access to a domain of representation that is beyond the regimes of either pictures or words — yet somehow in the shadow of both. [Rust 2008, 25]

The page is indeed one of the primary structural and compositional units of comics. In TEI the page is viewed as a "milestone," an empty marker within the flow of the text. In comics books (and similar documents), the page is not an arbitrary milestone, but a compositional feature — a composed container for "panels" and text.



Figure 2. The third of twelve miniatures from the *Electorum parvum seu Breviculum*, a compilation of the writings of philosopher, theologian, logician, and poet Ramon Llull (1232?-1316). [Llull and Pereira 1990, 12]



Figure 3. The ninth of twelve miniatures from the *Electorum parvum seu Breviculum*, a compilation of the writings of philosopher, theologian, logician, and poet Ramon Llull (1232?-1316). [Llull and Pereira 1990, 35]

The miniatures pictured in Figure 2 and Figure 3, tightly integrate text and image. Moreover, we find discreet, bordered, sequential images, similar to comic book panels, and text representing speech, like comic book word balloons, originating from the figures' mouths. In Figure 3, below the "panels" we find explanatory text, similar to the narrative captions found in comic books. Such examples illustrate formal and structural features of comics shared by many other document types. The analytical and descriptive strategies of TEI and CBML may prove useful in studying not just modern comics but an even more historically and culturally diverse assortment of document types.

16

Perhaps the most significant motivation behind the development of CBML is the wish to support for comic books and related documents the interpretive strategies and inscriptions of readings — discussed in the introduction above — that occur when a text is encoded. Unlike the traditional academic essay, in which a reading or analysis is sprinkled with relevant quotations from primary source materials, an encoded text may combine a digital representation of a full source text inscribed, in the form of markup, by a reading and performance of the text. Like the image and text of a comic book, the source text and the scholar's reading and analysis are inextricably intermingled in the encoded document. Encoded documents are metadocuments "which describe and enhance information but also serve as performative instruments.... As texts that describe a language, naming and articulating its structures, forms, and functions they [metalanguages and metatexts] seem to trump languages that are used merely for composition or expression" [Drucker 2009, 11]. CBML facilitates more transformative reading strategies that have developed in response to the increasing numbers of high-quality digital and digitized documents. Literary scholar Franco Moretti has popularized the term "distant reading" to describe strategies for analyzing massive numbers of literary and historical documents:

[I]f you want to look beyond the canon ... close reading will not do it. It's not designed to do it, it's designed to do the opposite. At bottom, it's a theological exercise — very solemn treatment of very few texts taken very seriously — whereas what we really need is a little pact with the devil: we know how to read texts, now let's learn how not to read them. Distant reading: where distance ... *is a condition of knowledge*: it allows you to focus on units that are much smaller or much larger than the text: devices, themes, tropes — or genres and systems. And if, between the very small and the very large, the text itself disappears, well, it is one of those cases when one can justifiably say, Less is more. If we want to understand the system in its entirety, we must accept losing something. [Moretti 2000, 57]

Digital technologies facilitate both close and distant reading strategies. Individual comic books and graphic narratives may be better understood through strategies of close reading, with careful attention paid to minute details of linguistic and visual language and document features, but comic books as a system and extended comic book narratives — which often unfold over decades in thousands or tens of thousands of individual documents by hundreds of creators (writers and artists) — benefit also from the distant reading (of text, image, and metadata) from comic book documents.

CBML aims to support such digitally-enabled and embedded performances and readings in comic books and across a range of graphics-intensive documents. This motivation is not unique to the study of comics or the development of CBML, but comic books and related documents do have particular, if not unique, features that must be supported in markup languages if we are to realize this goal.

Encoding the Comic Book

The following sections examine features and content of comic book documents and propose encoding strategies for representing, documenting, and analyzing those features and content types. These encoding strategies incorporate both new elements introduced in the CBML customization and existing TEI elements.

The comic book is a particularly complex class of document. A typical comic book, from the early twentieth century to the present, contains diverse content types. Alongside and intermingled with comics content, we may find prose fiction, advertisements, editorial and promotional content, fan mail, and bibliographic metadata about creators and publication details.

"Comics" Content: Sequential Art & Text

By "comics" content, I refer to the sequential art — usually combined with text — of the comics narrative. Comics content is typically divided into panels (usually with clearly delineated borders) that include images, narrative captions, word balloons, and sound effects (POW! SMASH! FOOM!).

18



Figure 4. Two pages from *The Fantastic Four* #51. This example contains many of the familiar elements of comics: panels, pictures, narrative captions, word balloons, sound effects, motion lines, etc. [Lee and Kirby 1966a]

A great deal of comics scholarship examines these individual elements in the context of the aesthetics, architectonics, and grammar of comics — how comics work, the systems and semantics, the structural components, the mutual dependencies of text and image, and how these textual and graphic components function together to convey information, narrative, plot, character, etc.^[10] Perhaps the most well-known and influential such study is Scott McCloud's *Understanding Comics*, a meta-comic about the art and form of comics, that includes this frequently-cited definition of comics:



Figure 5. From Scott McCloud's *Understanding Comics*, a page summarizing McCloud's definition of comics as "sequential art": His full definition reads: "Juxtaposed pictorial and other images in deliberate sequence, intended to convey information and/or to produce an aesthetic response in the viewer." [McCloud 1993, 9].

A possible objection to McCloud's definition is its subordination of text to image. In fact, McCloud's definition makes no mention of text. Certainly one can find examples of comics without any words, but, as McCloud acknowledges, the vast majority of comics do in fact contain a great deal of text of different types. The intricate and complex interplay between text and image is one of the defining features of most comic books. While we typically say that we *watch*, *view*, or see a film, we more often say that we *read* a comic. Lawrence L. Abbott, Henry John Pratt and others have emphasized the characterization of comics as a "read" medium [Abbott 1986] [Pratt 2009]. Abbott stresses this characterization in part to bolster his argument for the preeminence of text over image in comics: "The perceiver is, after all, termed a 'reader' — and the subordination of the pictorial to the literary in comic art is one of the distinctive text — the speech balloon — of comics: "The speech balloon is a defining element of the comic because it establishes a word/image unity that distinguishes comics from pictures illustrating a text" [Carrier 2000, 4]. The ability to identify and characterize textual components of comics art and to describe and analyze the interplay between text and image is one of the chief aims of CBML. TEI provides a host of options to support extensive analysis, textual criticism, and annotation. CBML suggests applications of existing TEI markup for encoding many of the distinctive features of comics and also provides a handful of additional elements for distinctive features that are not adequately handled by existing TEI markup.

In the discussion that follows, I shall use the terms "comics" to refer to sequential art (and text) and "comic book" to refer to a document that contain comics content and, optionally, other content types.

<cbml:panel>

The panel — encapsulating the constituent parts of image, text, and sound effects — is the primary building block of meaning in a comics text:

The panel is the fundamental unit of comic art.... The panel is the smallest unit in which the complex interaction of text and picture operates, and one notices quickly that the "text" in comic art takes form according to an elaborate series of conventions. [Abbott 1986, 156]

CBML provides the <cbml:panel> ^[11] element to represent this basic structural unit of comics. <cbml:panel> is a modification of TEI's <div> element, which represents a generic subdivision of the text in the TEI model.



Figure 6. The fifth panel of page 6, from *Captain America* #193 (January, 1976), edited, written, and drawn by Jack Kirby. [Kirby 1976]

```
<cbml:panel ana="#action-to-action" characters="#cap #anon man" n="5" xml:id="eg 000">
   <cbml:caption>
     Cap acts quickly to tranquilize the gun-happy pedestrian...
   </cbml:caption>
   <cbml:balloon type="speech" who="#cap" xml:id="eg 007">
      A little <emph rendition="#b">sleep</emph> will do wonders for you!
   </cbml:balloon>
   <sound> SPLAT! </sound>
   <cbml:balloon type="speech" who="#anon man"> Ugh! </cbml:balloon>
</cbml:panel>
```

Example 1. CBML fragment illustrating an encoded panel from Captain America #193. See Figure 6 above.

A few details of the above CBML fragment are worth noting, particularly the way in which both standard TEI elements 26 and attributes are used alongside custom CBML attributes to describe the structure and semantics of the panel.

The <cbml:caption> and <cbml:balloon> elements, described in more detail below, are used to encode these common components of comics panels. A TEI element <sound> is used to encode another common panel element, the graphical/textual representation of a sound in the narrative. One might imagine this element to be as peculiar to comics as the narrative caption or balloon, yet TEI provides the <sound> element, which seems to suffice. The TEI Guidelines explain that <sound> is intended to "describes a sound effect or musical sequence specified within a screen play or radio script" [TEI 2010a]. Of course a comic book is not a screen play or radio script, but comics and film share many similarities, and the scholarship on comics often notes these similarities.^[12] That the TEI element <sound>, designed to describe a phenomenon found in film, works so well, without any particularly wrenching semantic deformation, in the context of comic books is further evidence of the similarities shared by these two forms.

The @n attribute on <cbml:panel> is a TEI "global" attribute, meaning it is available on all TEI elements. The @n attribute may be used to provide a number or other label to the element. Depending on the needs of one's content or analysis, one could use @n to provide a descriptive label or sequential number to a panel. In the example above, n="5" indicates that this panel is sequentially the fifth panel among a larger group of panels, such as those appearing on a single page.^[13]

More interesting perhaps is the @characters attribute, a custom CBML attribute used to identify the characters 29 appearing in any given panel. The @character attribute contains, in TEI lingo, one or more data.pointer values. In the example above #cap and #anon man point to elements, with unique identifiers of cap and anon man, in the <teiHeader>. These elements in the <teiHeader> provide more detailed information about the characters Captain America and "anonymous man." The <teiHeader> is a mandatory and potentially very large TEI element that prefaces the encoded transcription of the document. The <teiHeader> provides a great deal of detailed descriptive information and other metadata about both the digital file and the source document. And it may, for instance, provide detailed information about all the characters appearing within a given document. With a large collection of comic books encoded using these mechanisms from CBML and TEI, a search interface could provide users the ability to retrieve instantly, from thousands or tens of thousands of comics, all the panels containing a character of interest, regardless of whether or not the character's name appears in the original panel.

In Example 1 above the TEI @ana attribute is used to describe the transition from the previous panel to the current panel. The attribute @ana, short for analysis, is another global TEI attribute (activated by inclusion of the optional TEI analysis module) that points to one or more elements containing interpretations of the element on which the @ana attribute appears. Typically @ana points to one or more <interp>, short for interpretation, elements. The document from which the above fragment is taken also contains this code:

30

27

```
<interpGrp resp="#jawalsh" type="panelTransition">
   <desc>
      The <gi>interp</gi> elements below include vocabulary
      that may be used to characterize the transition from the previous
      panel <emph>to</emph>the current panel. The transition
      types listed below are defined by Scott McCloud in his
      <bibl><title>Understanding Comics</title> <biblScope</pre>
      type="pp">(70-72)</biblScope></bibl>.
   </desc>
   <interp xml:id="moment-to-moment"/>
   <interp xml:id="action-to-action"/>
   <interp xml:id="subject-to-subject"/>
   <interp xml:id="scene-to-scene"/>
   <interp xml:id="aspect-to-aspect"/>
   <interp xml:id="non-sequitur"/>
</interpGrp>
```

Example 2. CBML fragment listing a vocabulary that may be used to characterize transitions from one panel to the next. The above example uses McCloud's transition types defined in his *Understanding Comics* [McCloud 1993, 70–72]. Of course, the encoder/editor may choose to augment or replace McCloud's vocabulary with any suitable typology.

The above code provides an "interpretation group" (<interpGrp>) of type "panelTransition." The @resp attribute points to a unique identifier of the agent responsible for the interpretations, for instance, the editor or author of the encoded document. A description (<desc>) inside <interpGrp>, explains the purpose and rationale for the interpretation group. Finally, a sequence of <interp> elements lists six possible panel-to-panel transitions, using a vocabulary borrowed from Scott McCloud, as explained in the description for the interpretation group. The @ana/<interp> mechanism is an extremely flexible one provided by TEI to encode multiple avenues of interpretations and analysis. One could, for instance, use this mechanism to record typologies related to visual motifs, actions and gestures, or gender representations and roles.

<div type="panelGrp">

Most theories of comics distinguish between a single-panel "cartoon" — such as Bill Kean's *The Family Circus*, Tom Wilson's *Ziggy*, or Gary Larson's *The Far Side* — and multi-panel comics, in which the sequence and juxtaposition of multiple panels plays a fundamental role in the production of meaning in the document. McCloud explains:



Figure 7. Scott McCloud distinguishes between comics and "cartoons":

Single panels are often lumped in with comics, yet there's no such thing as a sequence of one! Such single panels might be classified as "comic art" in the sense that they derive part of their visual vocabulary from comics — But ... they're no more comics than [a] still of Humphrey Bogart is film.... They are cartoons ... and there's a long-standing relationship between comics and cartoons. — But they are not the same thing! One is an approach to picture-making — While the other is a medium. [McCloud 1993, 20–21]

Panels, especially in print comic books, are not presented as a continuous, unbroken stream of narrative units, nor are they randomly laid out in sequence across a series of pages. The total collection of panels that make up a comics narrative are subdivided into panel groups. The sequence of panels that a reader has in her field of vision is typically an intentionally composed composite component of the larger work. For instance, a panel depicting a moment of high suspense might occur as the last panel of a recto page, forcing a lingering suspenseful pause while the reader turns the page to reveal the resolution of the suspense in the first panel of the verso page. These carefully composed panel groups often correspond to a physical page, but a panel group may share a physical page with advertisements or other content or may correspond to two physical pages of a two-page spread, in which panels are read from left to right, top to bottom across two opposing verso and recto pages. To capture this level of composition, CBML adopts the TEI <div> element, which is meant to encode a subdivision of the text. The @type attribute on <div> is used to indicate the type of textual subdivision, e.g., chapter, act, scene, canto, etc. In this case we use the string "panelGrp" to indicate the division is a grouping of panels. Another attribute, @subtype, is also available to distinguish further between different types of panel groups. Of course, the generic <div> element may also be used in CBML to encode other subdivisions,

such as chapters or "parts," which are commonly found in comics.

```
<div type="panelGrp" xml:id="eg_002">
   <cbml:panel characters="#david #samson">
      <cbml:balloon type="speech" who="#david">
         What a funny looking truck outside here ... Never saw one like
         it before!
      </cbml:balloon>
      <cbml:balloon type="speech" who="#samson">
         That's strange! What's it look like?
      </cbml:balloon>
   </cbml:panel>
   <cbml:panel characters="#samson #david">
      <cbml:balloon type="speech" who="#samson">
         You're right--I never saw one like this before!
      </cbml:balloon>
      <cbml:balloon type="speech" who="#david">
         Wonder what it's doing here?
      </cbml:balloon>
   </cbml:panel>
   <cbml:panel characters="#samson #david">
      <cbml:balloon type="speech" who="#samson">
         What the--!
      </cbml:balloon>
      <cbml:balloon type="speech" who="#david">
         Gas---Help!
      </cbml:balloon>
   </cbml:panel>
   <cbml:panel characters="#samson #david">
      <cbml:balloon type="speech" who="#samson">
         No time to look for doors now!
      </cbml:balloon>
      <sound>Crash!</sound>
      <fw place="lower-left" type="pageNum">1</fw>
   </cbml:panel>
</div>
```

Example 3. CBML fragment illustrating the use of <div type="panelGrp">. Panels are intentionally grouped and composed in larger compositional units, often corresponding to a physical page surface.

<cbml:caption> and <cbml:balloon>

A number of textual and graphic elements make up the content of individual panels. Textual content is found primarily in narrative captions and word balloons. Abbott provides a brief description of these elements:

Narration is placed in squared-off areas [narrative captions], usually colored yellow in comic books. Dialogue is located in white "balloons" that include little pointed projections indicating the speaker. Unspoken thoughts follow a similar pattern except that the balloons are billowy, and a string of small white circles leads to the thinker. Both narration and dialogue are recognized as extra-visual phenomena that may share space in the panel plane with the drawing but are not part of the scene. The visual assumption is that narration and dialogue lie on the plane of the opening through which one views the scene, augmenting the pictorial element but not part of it. [Abbott 1986, 156]

Abbott's description is accurate but incomplete. In addition to speech and thought balloons, one may find other types of balloons in comic book content. For instance, since so many genres of comics concern supernatural, speculative, fantastic, and science fiction narratives, "telepathic" balloons are not uncommon and are visually distinct from standard speech or thought balloons.



Figure 8. Professor Charles Xavier — leader of the X-Men, a team of super-powered mutants — is one of the more famous telepathic characters in comic books. The panel above shows Professor Xavier sending a telepathic message to his X-Men and provides an example of the distinctive balloons used to represent telepathic messages [Lee and Kirby 1963]. The balloon is styled similarly to a thought balloon but has additional clusters of small straight lines radiating out from the billowy, bubbly thought balloon. Such a balloon might be encoded as <cbml:balloon type="telepathic">type="telepathic"><

In many comics narratives, distinctly styled "audio" balloons may be found emanating from radios, televisions, telephones, walkie-talkies, hi-fi speakers, and other devices. These "audio" balloons are usually represented by jagged pointy borders, perhaps suggestive of the electricity that powers the audio source.



Figure 9. An example, from *The Amazing Spider-Man* #5 (October 1963), of an "audio" balloon, in which the sound/speech is produced by a television broadcast. Note the distinctive, jagged appearance of the balloon outline. The <code>@subtype</code> attribute is used to indicate that this is a television broadcast, as opposed to radio or some other broadcast medium. [Lee and Ditko 1963]

```
<cbml:panel characters="#spidey #jjj" n="3" xml:id="eg_ae1">
  <cbml:balloon rendition="#uc" type="audio" subtype="telecast" who="#jjj"
  xml:id="eg_006">
    My name is J. Jonah Jameson, publisher of <title rendition="#b">Now</title>
    magazine and the <title rendition="#b">Daily Bugle</title><mph
rendition="#b">!</emph>
    I am sponsoring this program in the public interest, to expose
    <emph rendition="#b">Spider-Man</emph> to the public as the menace he is!
    </cbml:balloon>
</cbml:panel>
```

Example 4. CBML code to represent the panel depicted in Figure 9.

One also finds a great deal of graphic stylistic variation among speech balloons. For instance, a character like the Vision, from Marvel Comics' *The Avengers* superhero team, is usually depicted with stylized speech balloons that highlight his android origins. Interestingly, when the Vision made his first appearance in *The Avengers* #57 in October 1968, his speech balloons were not styled differently from other characters. Later, the text inside the Vision's speech

balloons would be styled in a hand-written italics. The rectangular balloons with rounded corners were introduced in issue #91 in 1971. A yellowish tint was added to the balloons two months later in issue #93 and this convention has now persisted for over thirty-five years. This evolution of the graphic representation of the Vision's speech might be tied to an analysis of other aspects of the character's development.



Figure 10. The Vision's speech balloon is styled to suggest his android origins. [Busiek and Pérez 2000b]

The Vision's speech balloon is, however, still a speech balloon, but it is styled differently than other conventional speech balloons. TEI is equipped with a suite of elements and attributes (<rendition>, @rendition, and @rend) to describe the styling, or "rendition" of various elements in source documents. In primarily textual documents, the <rendition> element and related attributes might be used to describe details such as font family, font size, justification, and son on. The rendition features of TEI may be used in CBML contexts to describe graphical features, for instance to provide a detailed description of the distinctive styling of the android's speech balloon.

35

Another convention for presenting characters' speech and thoughts in comics eliminates balloons entirely and instead uses captions that incorporate both narrative text and the dialogue/monologue text usually found in balloons. At the other extreme are comics that contain balloons but lack narrative captions. The presence or absence of captions and balloons are particularly interesting examples of meaningful structural variation found in comics. The basic structures of the graphic panels and pictorial elements may be similar, but the structures of the textual elements are completely different. For McCloud, who privileges the pictorial, the different textual structures may be of interest but do not affect the content's status as comics. For other scholars who privilege the textual over the pictorial — Carrier, for instance — the lack of text balloons may mean that such documents are simply illustrated stories and not comics at all [Carrier 2000, 4, 27–45]. Certainly, these structures have important generic implications. The use of separate narrative captions and balloons results in structures that closely resembles the textual structures found in drama and film, while the absence of balloons results in structures that more closely resemble prose fiction. A complete absence of narrative captions is a likely indicator of the absence of a narrator in the text, while comics with a high number of narrative

captions might indicate a very strong narrative presence. An examination of the absence, presence, and frequency of narrative captions and balloons could be a useful strategy for analyzing the role of narration and the narrator in comics generally or in a particular title or author.



Figure 11. Three panel's from Gibbons and Sook's *Kamandi*. These panels illustrate a style of comic book that dispenses with balloons and includes all text — narration and dialogue — in narrative captions [Gibbons and Sook 2009]. Hal Foster's syndicated Sunday comic strip *Prince Valiant* also famously used this style of caption and is an obvious source for Gibbons and Sook.

In CBML, narrative captions are encoded using the <cbml:caption> element. TEI provides its own <caption> element, described in the *Guidelines* as "the text of a caption or other text displayed as part of a film script or screenplay." The examples provided in the *Guidelines* clearly indicate that TEI's <caption> element is intended for relatively brief and sparse captions, and the content model for the TEI <caption> element will not accommodate the complexity of many comics caption> element, CBML provides its own <cbml:caption> element, with a richer content model, in the CBML namespace. For captions like those in Figure 11, with embedded dialogue, one would use the TEI <said> element to demarcate the speech or thoughts inside <cbml:caption>.

Diegetic text and <floatingText type="diegetic">

Diegetic documents are another important textual and graphic feature of comics narratives. Diegetic documents appear in the comics art as part of the narrative's fictional universe. Such documents can be seen and "read" by the narrative's characters. We find diegetic documents or text appearing as street signs, store front signage, billboards, newspaper headlines, or even full newspaper articles. Pratt discusses clearly diegetic text and the degrees to which other text (balloons, captions, sound effects) may also be diegetic:

The words within comics have an interesting relationship to the diegesis, the story world that is "real" to and hence can be experienced by the characters who populate it. When a street number or a postcard is depicted in a comic, it is diegetic. Not only the reader, but also the characters can see it. However, only the reader can see the other types of words that occur in comics. The characters cannot see word balloons, sound effects, or narration (leaving aside cases where, for example, comics artists playfully have their characters interact with word balloons: breaking them, using them to float, and so on). But these features are not exactly non-diegetic. Though characters cannot see speech balloons, they can hear the words in them, and presumably each character is aware of the contents of his or her own thought balloons. When there are sound effects, characters can hear them, though the sounds heard within the diegesis may not be exactly the same as the sounds depicted in words. "Kablammo!" may be onomatopoetic, but it cannot capture the exact sound of an explosion. And characters may even be aware of narration, as is the case where one of the characters is also the narrator, or in the unusual situation where a character reacts directly to an impersonal narrator. [Pratt 2009, 108] 37

The term *diegetic* is frequently used in film studies to distinguish between the film score music that plays in the background but cannot be heard by the characters and the diegetic music that is part of the film's narrative sphere, for instance, when a character in the film plays the guitar or when music plays in a bar where characters meet. Famous examples from film include many of the scenes set in Rich Blaine's Café Américain in *Casablanca* and the music played by the "Cantina Band" in the Mos Eisley Cantina scene in the original 1977 *Star Wars* film. Below are examples of diegetic text in comics.



Figure 12. From *Amazing Fantasy* #15, Peter Parker walks by a poster to the science exhibit where he receives the radioactive spider bite that results in Parker's Spider-Man powers. The poster provides an example of diegetic text, text that is part of the comic's narrative environment and that may be "read" by the characters as well as the reader. [Lee and Ditko 1962]

```
<cbml:panel characters="#pparker" n="1">
   <cbml:balloon rendition="#uc" type="speech">
Some day I'll show them! <sound>sob</sound>
        Some day they'll be sorry! --Sorry that they
        laughed at me!
   </cbml:balloon>
   <floatingText subtype="poster" type="diegetic">
        <body>
            <head rendition="#uc #large">
                Science<lb/>
                Exhibit
            </head>
            Experiments<lb/><lb/>
                in<lb/>
                Radio-Activity
            Open<lb/>
                to the<lb/>
                Public
            <ab rendition="#right-arrow #red">
                Room 30
            </ab>
        </body>
   </floatingText>
</cbml:panel>
```

Example 5. TEI's <floatingText> element is used to encode diegetic text from the panel in Figure 12.



```
<div type="panelGrp" xml:base="eg/cbml eg.xml">
    <cbml:panel n="1">
        <cbml:balloon rendition="#uc" type="speech" who="#jjj"> Well, let's just see
with the <emph
                rendition="#b">distinguished competition</emph> has for their headline
this morning... </cbml:balloon>
        <floatingText subtype="newspaper" type="diegetic">
            <body>
                <ab type="masthead">
                    <title rendition="#uc #large #center"> New York Globe </title>
                    <1b/>
                    <hi rendition="#small #center #uc"> New York's Oldest Daily
Newspaper </hi>
                </ab>
                <div type="news_story">
                    <head rendition="#x-large #uc #b" type="headline"> Webbed
Wonder<lb/> Wows City!
                    </head>
                </div>
            </body>
        </floatingText>
   </cbml:panel>
    <cbml:panel n="2">
        <cbml:balloon rendition="#uc" type="speech" who="#jjj">
```

```
Huh.
            And let's see, what did the <title rendition="#b">Journal</title> run</title>
this
                morning?
        </cbml:balloon>
        <floatingText subtype="newspaper" type="diegetic">
            <body>
                <ab rendition="#center #large #uc" type="masthead"> The New York Journal
</ab>
                <div type="news story">
                    <head rendition="#x-large #uc #b #center" type="headline"> A Spider-
Man<lb/>
                        Among Us </head>
                </div>
            </body>
        </floatingText>
    </cbml:panel>
    <cbml:panel n="3">
        <cbml:balloon rendition="#uc" type="speech" who="#jjj">
            And...<emph rendition="#b">what,</emph> pray tell, did the <title</pre>
rendition="#b">Daily
                    Bugle</title> decide to run this morning?
            Some fat cat's <emph rendition="#b">house</emph> catches on fire.
        </cbml:balloon>
        <floatingText subtype="newspaper" type="diegetic">
            <body>
                <ab rendition="#center #large #uc" type="masthead"> Daily Bugle </ab>
                <div type="news_story">
                    <head rendition="#x-large #uc #b #center" type="headline">
Osborne<lb/>
                        Burning</head>
                </div>
            </body>
        </floatingText>
    </cbml:panel>
    <!--->
</div>
```

Example 6. TEI's <floatingText> element is used to encode diegetic text from the series of panels in Figure 13.

Prose Content: Fiction, Editorial, Promotional Material, Company and Industry News, Fan Mail

In addition to the comics content, a comic book will often contain more traditional prose of various sorts, including prose 38 supplements to the comics narrative; short stories, sometimes accompanied by illustrations;^[14] editorial and promotional content from publishers, editors, writers and artists; news items from the publisher; and fan mail, often accompanied by replies from the comic book creators.



UNDER THE HOOD

I.

THE HOOD

wit

<text><text><text><text><text><text><text><text><text><text><text>

Figure 14. From the first issue of Alan Moore's Watchmen, a comics art page juxtaposed against a page of more traditional prose fiction, an excerpt from the journal of Hollis Mason, a character in Moore's narrative. [Moore 1987]



Figure 15. Marvel's *Bullpen Bulletins* were a regular monthly feature in their line of comics. The *Bulletins* typically included "Stan Lee's Soapbox," a column by Marvel writer/editor/publisher Stan Lee, as well as smaller news items about new and upcoming comic books or television and radio programs featuring Marvel's characters or creators. In this installment of *Bullpen Bulletins*, Stan Lee announces the departure of artist Jack Kirby, Lee's longtime collaborator and co-creator of characters such as The Fantastic Four, Thor, X-Men, The Avengers, and more. This example is taken from *The Avengers* #79 (August 1970), although the same page would have appeared in Marvel's other comic book titles published during the same publication cycle. [Bullpen Bulletins 1970]

Comic book fan mail has great potential value to scholars studying comics. Fan mail highlights the social aspect of comic book creation and readership; the role of the reader in the creation, expansion, and elucidation of the larger, collective fictional universes of comics narratives; the often intense and intimate interaction between creator and audience; and frequent authorial or editorial commentary. Further, fan mail is often a source for important details such as the explicit identification of otherwise uncredited creators, including writers, artists, inkers, colorists, and letterers. There are many instances, as in Figure 16 below, in which individuals who would go on to become prominent figures in the comic book industry first appeared in comics as the authors of fan mail.

separate pieces of fan mail, highlighted in red, from individuals who would later become important figures in the comic book industry. The first letter is from Dave Cockrum [Cockrum 1964], the original artist on the enormously popular new X-Men team introduced by Cockrum and writer Len Wein in *Giant-Size X-Men* #1 (May 1975). The second highlighted letter is written by Roy Thomas, who became a writer and later Editor-in-Chief at Marvel Comics [Thomas 1964].

Advertisements

Print advertisements, particularly many of the types frequently found in comics, are themselves extremely complex documents. One cannot study the history of television in the twentieth century without careful attention to television advertising. Similarly, one cannot study the comic book without close consideration of the advertisements in document. In comic books from major publishers such as Marvel and DC, advertisements are, in terms of quantity, the dominant content type apart from the comics content.^[15] Of particular interest are the many advertisements that adopt conventions from comics. The intermingling of comics-based narrative with commercial advertisement becomes a key element in the production of meaning in the text. In addition to advertisements for third-party products, comics are rife with advertisements and other promotional materials for the comic books and other products of the publisher. These promotional materials contribute to the transformation of the publisher from commercial entity to a club or cult^[16] and to the process by which a reader becomes a "fan." Unfortunately, this important content is usually omitted from both print and digital "reprints" from the comic book publishers.

Advertisements are interesting on many levels. For instance, a scholar examining gender roles in comics may be interested to know that certain advertisements are addressed directly to boys ("Look, Fellows!") and that participation in certain activities promoted by the ads, such as selling *Grit* newspaper in one's neighborhood, required one to answer in the affirmative the question, "Are you a boy?" Later versions of the same ad would replace "Look, Fellows!" with "Look, Friends!" and "Are You a Boy?" with "Male or Female?"

40

Figure 17. An advertisement recruiting boys to sell *Grit* newspaper is addressed explicitly to boys, "Hey Fellows!," and the clip-out, mail-in coupon requires the applicant to answer the question "Are you a boy?" [Grit 1969]

Figure 18. Highlighted detail from Grit advertisement in Figure 17. [Grit 1969]

Williamsport, Pa. Send my first bundle	, 17701 of papers,	De selling helps a	nd prize offers.
Name	Sec. 17	- Taking	Sel - Star
Are You a Boy?	Age	Date Born	Year
In Care of		1000	Net Street
Street and No	W. TR	<u> </u>	_R.F.D
Post Office	Number		_State
State	Zip Co	de No	
Print Your Last Name Plainly Here		TTT	TIT

Figure 19. Highlighted detail from Grit advertisement in Figure 17. [Grit 1969]

Other advertisements found in comics are of interest because of the ways in which they appropriate the characteristics 42 of comics content, the dominant content type of the "host" document. One finds such appropriations increasingly in video and film as elaborate advertisements adopt the conventions of, for instance, music videos or short films. The televised broadcast of the National Football League's annual Super Bowl championship has become a showcase for such advertisements. Further intermingling of commerce and "art" is found in the relatively recent trend of advertisements in the movie theater and product placement within television and film. Comics anticipate such trends, on a mass scale, decades before they became commonplace in other media. Comic book documents encoded with CBML would allow one easily to identify, retrieve, and analyze advertisements that contain comic book elements such as panels and balloons.

The following examples illustrate some of the variety found in comic book advertisements and the extent to which the advertisements adopt the formal conventions of comics, from advertisements that include elements like word balloons to advertisements that contain miniature multi-panel comics narratives.

Figure 20. This advertisement for *Big Jim's P.A.C.K.* action figures incorporates comics-styled illustration, lettering and typography, captions, and word balloons. [Big Jim's P.A.C.K. 1976]

Figure 21. A 1968 advertisement for International Correspondence Schools (ICS) completely abandons traditional ad copy, format, and structure, instead employing the comics format to tell the fictional story of a high school dropout who becomes a success by using the advertised product. [International Correspondence Schools 1968]

<div type="advert">

The advertisements typically found in comic books may be adequately described using existing TEI elements, along with custom CBML elements for those advertisements that incorporate comics features such as panels and word balloons. The generic TEI <div> with its @type and @subtype attributes may be used as the container element for advertisements.

Figure 22. A small advertisement, from a page full of similarly sized advertisements, in Fantastic Four #65 (August 1967). The layout of the full page resembles the layout of a page of classified ads in a daily newspaper. [Crown 1967]

```
<div type="advert" xml:id="eg_003">
```

```
<head rend="background-color:black; color:white;" rendition="#x-large #center</pre>
#uc">Poems Wanted</head>
   <ab rendition="#center #uc" type="floatingHead">To Be Set to Music</ab>
   Send one or more of your best poems today for <emph rendition="#uc">free
examination</emph>.
        Any subject. Immediate Consideration.
   <ab rendition="#center #uc" type="floatingHead">Phonograph Records Made</ab>
    <ab>
        <address>
            <addrLine rendition="#uc #large #center">
                <orgName>Crown Music Co</orgName>
            </addrLine>
            <addrLine> 49 W. 32 St., Studio 11, New York 1 </addrLine>
        </address>
    </ab>
```

MS

</div>

Example 7. TEI/CBML code for the small advertisement featured in Figure 22.

Related Documents

Many other publications and content types that would not typically be classified as comic books incorporate the formal elements of comics. For instances, there are a number of publications that combine news, interviews, reviews, and

other journalistic features about comics and the comic book industry. These publications include fanzines, such as *Alter Ego*; official fan club publications, such as Marvel's *FOOM Magazine*; magazines such as *Wizard* or *The Jack Kirby Collector*; and Web sites, such as *Comic Book Resources*.

Figure 23. The cover (left) and an interior page (right) from Issue 2 (Summer 1973) of *FOOM Magazine*, the publication of Marvel's 1970s fan club FOOM [FOOM 1973, front cover, 11]. Many comics publishers or individual publications would provide readers with an opportunity to formalize their status as fans by joining official fan clubs, such as Marvel's Merry Marvel Marching Society and later F.O.O.M., or Friends of Ol' Marvel. These fan clubs also have publications, the digital representation of which can benefit from CBML encoding. These fan publications are important historical documents and include interviews with comics creators, news items, comic book checklists, games and puzzles, photographs and illustrations, previews, advertisements, and comics content in the form of covers and sample pages and panels from upcoming issues, reprinted comics, and original comics content created for the fan magazines.

The existence and relevance of these many comics-related publications is an important reason behind the decision to base CBML on TEI. Comic book content does not exist in isolation from other document types but is often integrated into essays, news articles, reviews, scholarly criticism, and so on. TEI provides robust mechanisms for encoding these more familiar document types, while CBML supplements TEI with features necessary for encoding the integrated comic book content.

One of the most noteworthy aspects of TEI is its many elements, attributes, and other structures for describing components of scholarly textual editions, including textual variants among multiple printings and editions, authorial manuscripts and typescripts, notes, annotations, and marginalia. Like many other modern documents, comic books can have complicated publication and production histories, with multiple editions and reprintings. Likewise, published comic books have their origins in authorial manuscripts and typescripts and original art with marginalia and notes of various types. See Figure 24 and Figure 25. The combination of CBML elements and existing TEI elements for describing manuscripts, typescripts, and textual variants provides a suitable suite of descriptors for encoding such documents.

46

SYNOPSES: THE FAITASTIC FOUR JULY 'GI SCHEDULE (F1) STORY \$1. INTRODUCTION. "WEET THE FANTASTIC FOUR"

This story is told in 2 chapters. Chapter one is 6 pages long. Chapter 2 is 5 pages.

This stary is tail in 2 support of the stary is tail in 2 support of the stary is tail in 2 support of the star is the star of the star of the star is the star of the star

East Bichards tolls duman and her brokher Johnny that He space ship is finally completed. He boyes to be first man to Kerse her be needs a pilot. Tay hirs her Grim. Sen is buy, early unpleased my why downed to want any part of project unit he seen Succes. He falls for Succes, and she manages to coar his into piloting ship. Bon is creaturjack pilot, st-war here, boot pilot semilable.

the four are about to begin flight, they are warmed against it by authorities. Told that me yet knows what effect commis rays will have on human bodies so far out in space. But feeting to go anyway. They fear that if they don't ye, lade any boat us to it.

At the rate the Communists are progressing in space, maybe we better make this a flight to the STRES, instead of just to Mare, because by the time this may goes on sole, the Runchnes may have already MUDE a flight to Marel) (NOTE: At

without elearance from the authorities, in the dead of night, they take off for the est star- very dramatically.

In space, on the way to the stars, FOGO2! They are bonharded by commite rays which penetrate the ship and which affect all four of the occupants. They can't continue the trip- have to turn back- are slucy to land alive. But they are all different now- they sense it- al-though they don't yet quite know HOW they've changed.

Sodemly, they can't see furmal bat they know sho's there: They can Hill har, They realize also has been invisible. She can not been wishin again. Later, she will buy a most the seen. Bar dothes of course can be seen, so it is only her fisch that is invisible. When ch takes her clothes off, abs's completely invisible. (I here this work's cern to cany in art weak. Bother walls to se show it, a complete and you have the complete and the second set of the second s

An for dehmap, Summ's hrother, whenever he gots excited, he bursts into flams. Seconds a Bumm forch, and can fly, an his body gots lighter than air. ST2 documt's last for zoro than 5 minutes. At and of five minutes, his flams evec out and he becomes normal gasim, with he gots andied again. His can't flams on for at least 5 minutes after he's gottes back to normal. Centes Associations told go he any prover burn augeou with flams, he ary enly burn reput. derive, dos-massy people. Add, he cannot team fireballs as the old Human Terch ends. States and the heart flams on the sense team fireballs as the old Human Terch

They think Röod Bichards, the pilot; is unaffected by comin rays, so be seens cornal-UNIL he tries to reach for scorthing. Then they ready in the STENTERD toward 'the thing be reached for. After saille they realize Bad's body has been bit BIERS. B, can get Admary, alongated, anything that you can do with rubber. We can equeese thru key-heles, etc. Of course, the more stretched-out he gote, the vector he getes- but the your remains that he can twick and stretch his body into alcost any shape. (So can even alter real to do all thing, so he can only maintain the strenge shaped for a very short period of time until the pain gets to be unbearable.

Finally, Ben Grimm stops out of the shadows. They all gasp-- his body has changed in the most grotesque way of all. Ho's cart of shapoless-- ho's become a THING. And, ho's groun more fantastically powerful than any other living hting. To is a stronger than an elephant. EUR, ho is so heavy that he boree very slouly-- ho's very ponderous, and there slow, pon-derous soveresta abould make his look very dividy-- more very ponderous, and there so low, pon-derous soveresta abould make his look very dividy-- nonto ther his appearance on the othere can, so he must war a coat with turnod-up coller, sunglamess, slouch hat, and glow when he goes out in public. Bu' take he hases 'en off, he is a THING'

Vanh he goes out in public. But when he takes 'en off, he is a fundit So such for who they are such how they got that way. Now, here's a gimnick I think we migh play up to advantage: Let's make The Thing the heavy- in other works, he's not really a good gay. Let's parts of the hearts Fantantic Four because they all got that way togother a they doids to remain a tens, and allos because he has a cruch on Suma-- but orbully, he is jacknow of Nr. Fantantic and dialkos Bunn Torch because they all got that way togother a fantantic Fautantic. Will Let's there the out of the start of the start is a start of the start kir Antantic Fautantic for the start of the start is a start in the start in being man kir. Fautantic for the start who he says interved in vinning fauna awy from kar kirs, because most of them how how heat their starts prover will last- or whether or the effect of the comic rays will one day-warr off theo).

us errect of the commin rays will can do not a surgery powers will list or whether of Anyway, the four of them decide to form a mult with the the them the the so that powers to which make them as they are and they think tay one that for to the the base power to and the source of the and THE THING, and you to possible that they can be the first of will ense which the normal forces of the verif cannot copy with Aci, to keep it all from getting to copy-goody, there is always frictions between Kr. National can first find, will human forch adding with Kr. J. Also, the other three are always aircaid of The filing getting to goody adding with Kr. J. Also, the other three are always aircaid of The filing getting to y to their control loss day and haring manifed with his manifes atrength. Occasionally also, to stop him. In other words, the Thing doem't have the other that the other three hours always be unpredictable.

So much for the introduction --- the preceding should have covered exactly 11 pages, consisting of 2 chapters. (Chapter one: 6 pages. Chapter 2: 5 pages)

The most two chapters, in which the Fantastic Four undertakes their first case, will also be undertakes for a initial total of D pages- fragmantemeters (3,5,5)

Figure 24. Stan Lee's original typescript synopsis of the origin story for Fantastic Four #1. Not all the elements from this initial conception, for instance Ben Grimm's romantic crush on Susan Storm, were implemented in the published story. [Lee 1961]

Figure 25. An example of original artwork for X-Men #17 [Kirby and Lee 1966]. Jack Kirby's marginal notes remain visible and provide valuable clues into the artist's contributions to the plotting and scripting of the story. [Kirby and Lee 1965?]

Linking Transcription and Image

The examples above illustrate how CBML can capture a transcription of the text found in a comic book. Other important features are also captured, such as the structure of the comic book; the sequence of pages and panels; the grouping of panels into compositional units that may or may not correspond to a physical page; classification of panel transition types and analysis of individual transitions from panel to panel. However, the examples above do not attempt to

describe the *pictures* one finds in the comic book, nor should they. Comic books are a visual, graphic art form combining text and image. CBML/TEI/XML is a text format. While one could certainly use CBML to describe details about any or all of the pictures in a comic book publication, such an effort would undermine the hybrid form of the comic book. The visual, pictorial, and graphic design elements of the comic book simply cannot be fully or adequately described or translated as text. While many design features common to textual documents, such as text size and font characteristics, may be reasonably and usefully described using common TEI techniques, it would be futile and impractical to attempt to describe every detail of every picture in a comic book document. The encoded document, with markup containing and describing metadata, structure, transcription, and analysis, should co-exist with and be linked to digital facsimile page images of the comic book. Thankfully, in current digital environments and on the Web, such linking and display of text and image is relatively simple, if time-consuming, to accomplish. Section 11.1 Digital Facsimiles in the *TEI Guidelines* describes existing TEI elements and attributes for linking digital transcriptions to digital facsimile images [TEI 2010c].

On the other hand, depending on the goals of any particular project employing CBML, some scholars will wish to add to the encoded text some description and analysis of the pictorial aspect of the comic book. If one has such needs, CBML and TEI provide suitable mechanisms. Typed TEI <note> elements may be used to provide detailed descriptions of the pictorial dimensions of the document. For example, one might use <note type="panelGrpDesc"> for a "panel group description" or <note type="panelDesc"> for a "panel description."

Figure 26. Two panels from *The Incredible Hulk* #114. Notice the common feature of a word balloon that begins in the first panel but crosses the gutter separating the panels and extends graphically into the second panel. Such features pose challenges for the strictly hierarchical structure of XML. [Lee and Trimpe 1969]

```
<div type="panelGrp">
    <!-- Preceding panels omitted to conserve space. -->
    <cbml:panel characters="#hulk" n="7">
        <note resp="#jawalsh" type="panelDesc"> The hulk kneels in the grass, surrounded
by trees.
            He leans back with hands on his head, lamenting his transformation back into
Banner. </note>
        <cbml:balloon rendition="#jaggies" type="speech"> No!<lb/>
            <emph>No!</emph>
        </cbml:balloon>
        <cbml:balloon type="speech"> It feels like-- <lb/> my <emph>blood</emph>--<lb/>
is on
                <emph>fire!</emph>
        </cbml:balloon>
        <cbml:balloon type="speech">
            My <emph>skins--<lb/> shrinking--<lb/> shrinking--!!</emph>
             Can't--<1b/>
                <emph>fight</emph> it-- anymore--! 
        </cbml:balloon>
    </cbml:panel>
    <cbml:panel characters="#hulk #banner" n="8">
        <note resp="#jawalsh" type="panelDesc"> Set against an abstract background, a
sequence of
            four views of the Hulk's and Banner's head depicts the Hulk's transition
back to Bruce
            Banner. </note>
        <cbml:caption> Thus, there and then--<lb/> suddenly--uncontrollably--<lb/> the
awesome
                <emph>transformation</emph><lb/> starts--until-- </cbml:caption>
   </cbml:panel>
    <!-- <cbml:balloon -->
    <cbml:panel>
        <cbml:balloon type="speech" who="#cap" xml:id="eg_008"> A little <emph</pre>
rendition="#b"
                >sleep</emph> will do wonders for you! </cbml:balloon>
    </cbml:panel>
    <!-- captions -->
    <cbml:panel>
        <cbml:caption rendition="#uc" xml:id="eg 009"> Thus, there and then--<lb/>
            suddenly--uncontrollably--<lb/> the awesome <emph>transformation</emph><lb/>
            starts--until-- </cbml:caption>
    </cbml:panel>
</div>
```

Example 8. <note type="panelDesc"> is used in this code example to describe the pictorial content of individual panels. Similarly <note type="panelGrpDesc"> could be used to provide a description of the layout or other visual details of group of panels.

Another option is to use the @ana attribute, illustrated above, to apply analytical or interpretive schemes to the pictorial, graphic, and visual dimensions of the text.

Even when focusing solely on the text of comics, the transformation from handwritten or printed text to unformatted digital transcription may lose important information-bearing graphic and design features. For instance, in the examples in Figure 27 below, distinct fonts and balloons are used for particular characters in an attempt to represent graphically the tone, timbre, or other qualities of speech, resulting in a sort of visual or synaesthetic onomatopoeia. In such cases, TEI's <rendition> element and @rendition and @rend attributes may be used to describe such design features.

Figure 27. In this panel fragment from Avengers #27 (Apr. 2000), a great deal of information is being conveyed in the stylization of the speech balloons [Busiek and Pérez 2000a]. In particular, notice the distinctive font used for the dialog of Thor, the large figure in the center of the panel. This use of distinctive fonts for specific characters is also an indicator of changing modes of production for comics. Traditionally, comics were lettered by hand. It is now increasingly common, especially among the major comic book publishers, for the lettering to be done in desktop publishing and design software using digital fonts, many of which are designed to resemble traditional hand-drawn comic book lettering, providing a digital illusion of handcrafted script.

Challenges: Visual Complexity

The basic structures of CBML were based on representative samples of common and relatively simple structures and panel layouts. For instance, panels normally are arranged in fairly regular grid patterns. 2 x 3 panel grids (three rows of two rectangular, nearly square panels) are very common; 2 x 2 and 3 x 3 layouts are also prevalent.

Figure 28. A page from the Captain America feature in *Tales of Suspense* #84 (Dec. 1966) by Stan Lee and Jack Kirby illustrates a very regular 2 x 3 panel grid. [Lee and Kirby 1966b]

But variation from strict and regular grid layouts is ubiquitous. Panels of every size and shape — circles, triangles, irregular polygons, and fluid organic shapes — are also common, as are borderless panels with ambiguous boundaries.

Figure 29. Two Golden Age examples of irregularly shaped panels. On the left is a page by Carl Burgos from *Marvel Mystery Comics* (1941) [Burgos 1941]. On the right is a page by Sheldon "Shelly" Moldoff from *Flash Comics* #5 (1940) [Moldoff 1940].

Simplistic approaches to encoding the spatial and sequential relationships among such panels are foiled by frequent variations and complexities in panel size, shape, and arrangement and particularly by ambiguous sequential positioning of panels. As determination of meaning in a textual document is dependent on a particular, usually obvious, sequence of words, so is determination of meaning in a comic book dependent on a particular sequence of images and panels. While a meaningful panel sequence in comics is usually apparent, the conventions of graphic expression in comics are not as fixed, well established, or as well known by general readers as the conventions for textual expression and narrative. And just as writers and poets break rules and conventions, so do comics writers and artists play with the rules and conventions of comics narrative. In Figure 30 below, a single obvious sequence of panels is not at all clear — multiple valid sequential readings are possible. Also note the playful use of diegetic text, with bibliographic information about the creators and title of the story embedded in the pictures. Near the upper left of the page is a blue-tinted panel with the names of the creators of the story: "LEE STERANKO SINNOT ROSEN in Another Epic!" The title of the story, "Tomorrow You Live Tonight I Die!," appears on a calling card in the panel at the lower right-hand corner of the page.

Figure 30. A page by Jim Steranko and Stan Lee from *Captain America* #111 [Lee and Steranko 1969]. The page illustrates ambiguous panel sequence and playful use of diegetic text.

A <note> and omission of sequential numbering on the <cbml:panel> elements could be used in the encoding to address such sequential ambiguity.

Irregular and unusual panel arrangement, composition, and presentation may pose challenges for "reading" of the visual and textual language of comic books, just as difficult syntax, unfamiliar vocabulary, neologisms, or unconventional typography and punctuation may present challenges for reading primarily textual content. Most comics have panels with clear borders and generous gutters between borders. The borders and gutters between panels correspond to the punctuation and spacing used to enhance clarity and readability in written and printed texts. Panels without clear borders and gutters introduce another type of visual ambiguity in comics. Figure 31 illustrates a panel group with discreet pictorial moments that correspond to panels, but these "panels" lack conventional borders and gutters.

Figure 31. One of many interesting examples from Will Eisner's influential *Comics & Sequential Art*, this page is composed as one blended graphic lacking common panel borders and gutters. Nonetheless, the larger image may be broken down into smaller sequential visual and textual gestures that correspond closely to conventional comic book panels. On the left is the original page; on the right, I have superimposed boxes over the image to identify the discreet moments or "panels" present in the larger composite image. [Eisner 1985, 75]

At least six such pictorial moments or "panels" may be identified in Eisner's page from Figure 31. The top of the image depicts a man and woman running through Charles de Gaulle Airport; the next element shows, from a distance, the couple as they ride up an escalator. The lines of the escalator tube, extending more or less horizontally across the entire page, cause this image to serve simultaneously as a discreet pictorial moment *and* as a border or gutter separating the "panels" above and below the escalator scene. Two more "panels" are found directly below the escalator. On the left, the couple rush through the crowded airport; on the right, they arrive at their seats on the plane they have been rushing to catch. A very subtle "panel" emerges quietly below these two panels as the aircraft, seen from a distance, takes off through the clouds. Like the distant view of the escalators, this image does double duty as "panel" and as border/gutter. In the final "panel" of the page, the man and woman, reclined in their seats, relax and resume their conversation. In those images (the escalator and the ascending aeroplane) that function as both panel and border/gutter, we might say, borrowing terminology from textual criticism, that the substantive content has merged with the accidentals of the document. By encoding these graphic and textual moments within <cbml:panel> elements, the scholar encoder has imposed an interpretation and asserted various claims about the document. Additional @ana attributes might be used to indicate those "panels" that function as both panel and border/gutter. <note> elements might also be used to discuss other issues related to the general graphic ambiguity of the page.

In Figure 26 above, we saw an example of a speech balloon crossing the gutter from one panel into another. Another type of visual ambiguity is introduced when pictorial elements cross panel boundaries and co-exist is more than one panel. Since different panels often depict different physical spaces and different moments in time, the presence of a pictorial element in more than one panel can be particularly jarring and can be used to establish complex spatial and temporal relationships. Figure 32 below shows an example of this type of visual ambiguity in an early WWII-era comic book:

Figure 32. A page from Simon and Kirby's *Captain America Comics* #1 from 1941. The page illustrates complex panel composition with pictorial and graphic elements crossing the gutters separating panels. [Simon and Kirby 1965, 168]

The panels above are not uniform in shape or size. Pictorial elements, such as the gun in the fifth panel and the purplesuited figure and motion lines in the bottom two panels, cross the gutter separating the panels and co-exist in multiple panels. These graphic moves suggest interesting spatial and temporal juxtapositions and facilite visual transitions from panel to panel, breaking down the clear separation of narrative moments and instigating a flow approaching (though still very far removed from) the rapid frame-to-frame transitions found in film.

Figure 33 below, from Morrison and Jone's *Marvel Boy* [Morrison 2000], shows a more recent example of a page with both ambiguous panel boundaries and overlapping panels. The central action — of the blue-cloaked golden figure lifting Marvel Boy into the air — is witness to its own genesis, visually encroaching on the panel depicting some moments prior in time. As Marvel Boy is lifted up into the underlying panel he becomes a spectator — with the reader — of himself within the previous moment. Likewise the blue-cloaked figure plants a golden boot upon a future moment that proceeds from the central action. An encoded CBML document might identify and analyze such transitions. One could develop a taxonomy to classify the sorts of overlap phenomena found in Figure 32 and Figure 33 and use the TEl <interGrp> and <interp> elements and the @ana attribute, discussed above, to describe and interpret the graphic features and the resulting spatial and temporal relationships.

Figure 33. In Grant Morrison and J.G. Jone's *Marvel Boy*, borderless and overlapping panels confuse and complicate temporal and spatial relationships. [Morrison 2000]

Serialization and Bibliographic Complexity

As is the case in many other art forms (fiction, television, film), serial publication of an ongoing narrative introduces a great deal of bibliographic and metadata complexity. While these issues are not unique to comics, I would argue that the degree of complexity is distinct in comics. Comic books typically contain serialized narratives. Often one finds "standalone" issues, which contain a complete story, and often stories extend over many issues, or even many issues of two or more serial titles.^[17] In the case of mainstream super hero comics, regardless of whether any particular story is more or less self-contained within a single issue or extends over many issues, these narratives exist in the context of larger narratives developed over many decades by dozens or hundreds of individual writers and artists. Rather uniquely, characters like Superman, Batman, Wonder Woman, Spider-Man, and the Hulk exist in ongoing, decades-long, crossgenerational narratives created by collaborative teams of writers, artists, and editors. Although we find similar issues of long-running narratives with changing collaborative teams in film and television, none approaches the longevity and frequency of production found in the comic book narratives.^[18] Most long-running comic book features, from Fawcett's original Captain Marvel to Spider-Man to Archie, have complex bibliographic histories involving networks of changing titles, creators, even publishers. These complex bibliographic histories, with multiple creators, are relevant to textual theories about the nature of authorship and the social and corporate nature of document production. In comic books we have examples of collaboratively-produced narratives coupled with collaboratively-produced visual art. The many metadata structures available in TEI lend themselves to description of these complex bibliographic histories and relationships.

The evolution of characters over many decades in serialized tales poses documentary complications that may be addressed through encoding strategies. For instance, one may be interested in the DC super hero known as the Flash. The general concept of a super hero named the Flash, with the power of super speed, published by the comic book publisher DC, is some sort (or some sorts) of bibliographic entity — a character, a subject, a title. But there is not one Flash character (or title) in the DC imaginative universe. The first Flash, with the secret identity of Jay Garrick, was created by Gardner Fox and artist Harry Lampert in *Flash Comics* #1 (January 1940). This character was revived in the 1950s, with a new costume, new origin story, and new secret identity — police scientist Barry Allen. Allen "died" in the 1985 series *Crisis on Infinite Earths* and remained "dead" for over twenty (real — not fictional) years . During Allen's absence, his similarly super-powered sidekick Wally West took on the persona of the Flash, as did Bart Allen, Barry Allen's grandson. Many other characters have undergone similar evolutions and transformations.

58

60

61

TEI is well-equipped to capture these complex details of character and identity. The prosopographical features of TEI provide a framework for modeling relations among both real people (e.g., comics writers, artists, and other creators) and fictional characters.

Conclusions and Future Work

Like most digital humanities projects and markup language development, CBML is an ongoing effort. Although many diverse examples have been shown here, the extent of textual and graphic variation found in the world of comics is astounding. Examples here have focused on superhero comics, the most enduringly popular genre in the history of the art form. However, many other important genres exist — romance, war, horror, science fiction, and fantasy comics — each with their own traditions and conventions. Increasingly, more realistic narratives, autobiographical works, and non-fiction works are presented as comics. Japanese manga and manga-influenced comics are very popular and possess a style and conventions distinct from western comics. Certainly many more difficult and challenging comic book documents will be found to test the general applicability of CBML and suggest additions and modifications to the markup language.

An important issue, not yet addressed in this discussion, is the issue of copyright. Modern comic books, which developed in the 1930s and since, are largely still under copyright protection. The most popular of these, those featuring the famous super heroes, are owned by massive media conglomerates. DC Comics (publishers of titles featuring Superman, Batman, Wonder Woman, the Flash, etc.) is owned by Time Warner. Marvel Comics (publishers of titles featuring Spider-Man, the Fantastic Four, Iron Man, the Hulk, etc.) was recently acquired by the Walt Disney Company. These media companies have reputations for jealously and fiercely protecting their copyrights. Copyright is doubtless a challenge for scholars studying comic books, just as copyright is a challenge for scholars studying film and many other twentieth- and twenty-first century works. But barring further restrictive copyright legislation, many of the older comic book publications will be entering the public domain in coming decades. Under current U.S. Copyright Law, Action Comics #1 (June 30, 1938), featuring the first appearance of Superman, would enter the public domain in 2033, 95 years after publication. And many comic books from the same period are already in the public domain because the copyright was not renewed, when such renewal was still necessary for an extension of copyright protection. In 2008, publisher Image Comics began their Next Issue Project, a series of new comics based on the narratives and characters from comic books published in the 1930s and 40s that had fallen into the public domain. The first Next Issue Project comic book was Fantastic Comics #24 [Fantastic Comics 2009]. It draws on stories and characters from Fox Feature Syndicate's Fantastic Comics, the final issue of which was #23, published in November 1941. The Next Issue Project has identified a number of other public domain titles, including Lev Gleason Publications' Silver Streak Comics (twentythree issues), Quality Comics' Crack Comics (sixty-two issues), and Harvey Comics' Speed Comics (forty-four issues). These four titles alone provide over 150 comic books that could be used for further collaborative development of CBML and published on the Web and in other media. Certainly there is a great deal more comics content in the public domain or for which permissions may be secured, and nothing prevents a scholar from applying CBML markup to any text as part of a strategy for reading, interpretation, and analysis. The end goal of markup is not and should not always be publication of a digital surrogate. The encoding of a text may be a rigorous intellectual activity that has great value as process, not just as product.

CBML, as described here, focuses primarily on overall structure of the document, textual content, and metadata. Significant work remains to be done giving similar attention to the pictorial dimensions of the comic book. I have argued above that CBML should not attempt to go too far in description of individual images and pictorial details, relying instead on the presence of facsimile page images. Nonetheless, in order to analyze the visual grammar and conventions of comics, additional visual and pictorial features will need to be "identified explicitly in order to facilitate processing by computer programs" [TEI 2010c]. Future work on CBML will include modeling frameworks for analysis of pictorial and graphic features and developing taxonomies for identifying such features, beyond the basic structural components of panels, balloons, captions, and sound effects.

62

63

64

65

Comic books are an endlessly fascinating medium and one from which the digital humanities can learn a great deal. The structural complexities of the content rivals or surpasses the complexity of content famously treated by digital humanities scholarship: fragmentary classical texts; medieval manuscripts; Blake's illuminated work; and Rossetti's double works. The combination of text and pictures and the role of visual communication in comics prefigures the multimedia and new media of digital environments and the prominence of pictures, graphics, and icons found on the Web and other modern digital interfaces. The corporate and collective authorship of many long-running comic book narratives and the interactive relationship between creators and readers/viewers/consumers of created content provide models of authorship and readership that are relevant to the collaboratively created digital content that we make and study in the digital humanities.

Ongoing development of CBML is documented at http://www.cbml.org/. One may download the most recent schemas and TEI ODD files (which define and document the CBML customizations of TEI) and example CBML instance documents. CBML-L, an email list for discussion of CBML, is available at https://iulist.indiana.edu/sympa/info/cbml-l. In addition to ongoing development and refinement of the CBML markup scheme, future research will involve the creation of a large corpus of pubic domain comics encoded in CBML and the development of interfaces for representing, exploring, and manipulating CBML documents and associated facsimile page images.

Acknowledgments

I would like to thank *DHQ* Editor Julia Flanders and the reviewers for their very valuable feedback on this article. I would also like to thank my colleagues Paul Aarstad and Michelle Dalmau for their careful reading and insightful comments.

Notes

[1] XML, or eXtensible Markup Language, is a widely-adopted metalanguage that specifies a set of rules for encoding documents and data and for creating application- and domain-specific markup languages for encoding documents and data. XML is a recommendation of the World Wide Web Consortium. See [Bray 2008].

[2] I share with many comics creators and scholars a dissatisfaction with the term *graphic novel*, finding it unnecessary, misleading and perpetuating of a false distinction. Wikipedia's article on Graphic novel includes a useful summary of many of the criticisms of the term.

[3] An example is Peter David's *Mascot to the Rescue!* [David 2008], a superhero-themed children's novel that integrates comics content with more traditional narrative prose. Brian Selznick's *The Invention of Hugo Cabret* [Selznick 2007] is an illustrated novel, with many of the illustrations subdivided into juxtaposed images reminiscent of comics panels.

[4] Readers from the scholarly markup and digital humanities communities will be familiar with many of the general issues about text encoding discussed here. I am also hopeful that this essay will attract readers from the comics scholarship community who may not be as familiar with text encoding, and so I go into more detail about general issues of text encoding than I might otherwise.

[5] The reader wishing to test whether or not it is indeed "plausible to imagine discussions over whether a given way of marking up a text is correct or incorrect" is invited to browse the TEI Mailing List (TEI-L) Archive, where she will find many lively and energetic debates on markup practices.

[6] Hundreds of scholarly projects are based upon underlying TEI-encoded texts and data. Examples includes my own projects: *The Algernon Charles Swinburne Project* and *Chymistry of Isaac Newton*, a collaboration with William R. Newman, Professor of History of Science at Indiana University. While not an exhaustive list, many other TEI-based projects may be found at http://www.tei-c.org/Activities/Projects/.

[7] See chapter 23 "Using the TEI" of the TEI Guidelines [TEI 2010b] and Roma, an online tool for "generating validators for the TEI."

[8] Additional examples of recent comic book film adaptations include *From Hell* (2001), *Ghost World* (2001), *Road to Perdition* (2002), *American Splendor* (2003), *Daredevil* (2003), *Hulk* (2003), *The League of Extraordinary Gentlemen* (2003), *Catwoman* (2004), *Hellboy* (2004), *The Punisher* (2004), *Batman Begins* (2005), *Fantastic Four* (2005), *A History of Violence* (2005), *Sin City* (2005), *V for Vendetta* (2005), *300* (2006), *Superman Returns* (2006), *4: Rise of the Silver Surfer* (2007), *30 Days of Night* (2007), *Ghost Rider* (2007), *The Dark Knight* (2008), *Hellboy II: The Golden Army* (2008), *The Incredible Hulk* (2008), *Iron Man* (2008), *Punisher: War Zone* (2008), *The Spirit* (2008), *Watchmen* (2009), *Whiteout* (2009), *Iron Man* 2 (2010), *Kick-Ass* (2010), *Captain America: The First Avenger* (2011), *Green Lantern* (2011), *X-Men: First Class, Thor* (2011), *The Dark Knight Rises* (2012), and *The Amazing Spider-Man* (2012). Many more film adaptations of comics are planned or in development. Extensive lists of films based on comics may be found on Wikipedia's " List of Films Based on Comics".

[9] The University Press of Mississippi publishes dozens of comics related titles. See, for instance, [Di Liddo 2009], [Gordon, Jancovich, and McAllister 2007], [Hatfield 2005], [Jeet 2008], [Kunzle 2007], and [Lee and McLaughlin 2007]. Other scholarly presses are also publishing monographs on comics studies. See [Aldama 2009], [Carlin, Karasik, and Walker 2005], [Carrier 2000], and [McLain 2009].

[10] Comics scholarship also includes historical studies, studies of individual writers and artists, and investigations of traditional literary issues of theme, plot, narrative and character; as well as representations of politics, class, race, and gender [Capitanio 2010] [Huebner 2009] [Wanzo 2009] [Coogan 2006] [Emad 2006] [Wright 2001]. Other fields, such as education and rhetoric, study the use of comics in the classroom to teach reading and writing [Jacobs 2007] [Ranker 2007] [Norton 2003]. Within the field of library science one finds research on the history of comics within libraries and on the development of comics and graphic novel collections [Matz 2004] [Highsmith 1993] [Hoffmann 1988].

[11] In this essay, XML elements are enclosed in angle brackets (<>) and displayed in a monospaced font. XML attributes, by common convention, are prefixed with the "at" symbol (@) and are also displayed in a monospaced font. The prefix cbml: identifies CBML elements added to TEI using TEI customization mechanisms, e.g., <cbml:panel> or <cbml:balloon>. Elements without the cbml: prefix are stock TEI elements, e.g., <div> or <epigraph>.

[12] See for instance Johanna Drucker's "What is Graphic about Graphic Novels?"" [Drucker 2008].

[13] Additional global attributes include <xml:id>, for providing an element with a unique identifier; @rend and @rendition, two related attributes used to describe how the element is rendered or styled in the source document; and @xml:lang, for indicating the language of the text in the element. More global attributes, such as those used for linking and analysis are added by including optional TEI modules in one's schema.

[14] Golden Age comics often included such prose narratives. In fact, the first published comics work of Stan Lee, the famous co-creator of such characters as Spider-Man, the Fantastic Four, the Hulk, and Iron Man, was a prose story in *Captain America Comics* #3 [Lee 1941]. More recently, Alan Moore frequently includes prose material that supplements the comics-based narrative. Examples may be found in Moore's *Watchmen* and *The League of Extraordinary Gentlemen*.

[15] As a random example, *Fantastic Four* #61 (April, 1967) contains thirty-six pages, including front and back covers, recto and verso. Of these thirty-six pages, twenty-one are devoted to comics content, twelve to advertisements, two to fan mail, and one page to other editorial content.

[16] Stan Lee, who was both a major creative force at Marvel Comics as well as its most effective and visible promoter, frequently addresses Marvel readers/fans as "True Believers."

[17] For example, the Marvel Comics "Avengers Defenders War" storyline from 1973-74 takes place in two serial publications, *The Avengers* (issues #115 to #118) and *The Defenders* (issues #8 to #11).

[18] The closest analog might be daytime soap operas, many of which have been in production for decades and are produced more frequently (with daily episodes Monday through Friday) than typical comic book issues are published. But the longest-running soap opera, *Guiding Light*, which started on radio in 1937, ceased production in 2009. The Superman and Batman narratives, which started in 1938 and 1939, respectively, continue to be told in multiple serial publications and of course in film and other media. Another analog might be the repeated refashioning of mythological, biblical, and Arthurian narratives throughout literary history, but such recasting of shared imaginative universes are more decentralized and lack the corporate control and continuity found in the ongoing comic book narratives or soap operas.

Works Cited

Abbott 1986 Abbott, Lawrence L. "Comic Art: Characteristics and Potentialities of a Narrative Medium". Journal of Popular

Culture 19 (1986), pp. 155-176.

- Aldama 2009 Aldama, Frederick Luis. Your Brain on Latino Comics: From Gus Arriola to Los Bros Hernandez. Austin: U of Texas P, 2009.
- Barnes 2012 Barnes, Brooks. "Avengers' Vanquish Box-Office Rivals". New York Times (May 7 2012), pp. C1-C1.
- Bendis and Bagley 2001 Bendis, Brian Michael, and Mark Bagley. *Ultimate Spider-Man #6. Ultimate Spider-Man*. Marvel Comics, April 2001.
- Big Jim's P.A.C.K. 1976 Mattell. "Big Jim's P.A.C.K". In Captain America #193. Marvel Comics, January 1976. pp. 9-9.
- Bray 2008 Bray, Tim, Jean Paoli, C.M. Sperberg-McQueen, Eve Maler and François Yergeau, eds. *Extensible Markup Language (XML) 1.0*. World Wide Web Consortium, 2008. http://www.w3.org/TR/REC-xml/.
- Bullpen Bulletins 1970 Marvel Comics. "Bullpen Bulletins". In The Avengers #79. Marvel Comics, August 1973. pp. 30-30.
- Burgos 1941 Burgos, Carl. "Marvel Mystery Comics". In Jules Feiffer, ed., *The Great Comic Book Heroes*. New York: Bonanza Books, 1965. pp. 84-98.
- Busiek and Pérez 2000a Busiek, Kurt, and George Pérez. The Avengers #27. The Avengers. Marvel Comics, April 2000.
- Busiek and Pérez 2000b Busiek, Kurt, and George Pérez. *The Avengers* #32. *The Avengers*. Marvel Comics, November 2000.
- **Capitanio 2010** Capitanio, Adam. "The Jekyll and Hyde of the Atomic Age: The Incredible Hulk as the Ambiguous Embodiment of Nuclear Power". *The Journal of Popular Culture* 43: 2 (2010), pp. 249-270.
- Carlin, Karasik, and Walker 2005 Carlin, John, Paul Karasik and Brian Walker, eds. *Masters of American Comics*. New Haven, CT: Yale UP, 2005.
- Carrier 2000 Carrier, David. The Aesthetics of Comics. University Park, PA: U of Pennsylvania P, 2000.
- **Cockrum 1964** Cockrum, Dave. "Letter to Stan Lee and Jack Kirby from Dave Cockrum". *Fantastic 4 Fan Page.* In *The Fantastic Four #22*. Marvel Comics, January 1964. pp. 29-29.
- **Conway 1976** Conway, Gerry, and Ross Andru. *Superman vs. the Amazing Spider-Man.* New York: DC Comics & Marvel Comics, 1976.
- Coogan 2006 Coogan, Peter. Superhero: The Secret Origin of a Genre. Austin: MonkeyBrain Books, 2006.
- Crown 1967 Crown Music Co. "Advertisement". In Fantastic Four #65. Marvel Comics, August 1967. pp. 22-22.
- David 2008 David, Peter. Mascot to the Rescue! New York: Harper Collins, 2008.
- **Di Liddo 2009** Di Liddo, Annalisa. *Alan Moore: Comics as Performance, Fiction as Scalpel*. Jackson: UP of Mississippi, 2009.
- **Drucker 2008** Drucker, Johanna. "What is Graphic about Graphic Novels?". *English Language Notes* 46: 2 (2008), pp. 39-55.
- **Drucker 2009** Drucker, Johanna. SpecLab: Digital Aesthetics and Projects in Speculative Computing. Chicago: University of Chicago Press, 2009.
- Eisner 1985 Eisner, Will. Comics & Sequential Art. Tamarac, FL: Poorhouse Press, 1985.
- **Emad 2006** Emad, Mitra C. "Reading Wonder Woman's Body: Mythologies of Gender and Nation". *The Journal of Popular Culture* 39: 6 (2006), pp. 954-984.
- FOOM 1973 FOOM Summer 1973. Marvel Comics Group. Print.
- Fantastic Comics 2009 Various authors. Fantastic Comics #24. Fantastic Comics. Image Comics, 2009.
- Flanders 2005 Flanders, Julia. *Digital Humanities and the Politics of Scholarly Work*. Thesis, Brown University: 2005. http://dev.stg.brown.edu/staff/Julia_Flanders/pubs/flanders_dissertation.xhtml"/.
- Gibbons and Sook 2009 Gibbons, Dave, and Ryan Sook. "Kamandi: The Last Boy on Earth". In *Wednesday Comics* #2. DC Comics, September 2009.

- Gordon, Jancovich, and McAllister 2007 Gordon, Ian, Mark Jancovich and Matthew P. McAllister, eds. *Film and Comic Books*. Jackson: UP of Mississippi, 2007.
- Grit 1969 Grit Publishing Company. ". Advertisement for Grit". In *The Fantastic Four* #93. Marvel Comics, December 1969. pp. 9-9.
- Hatfield 2005 Hatfield, Charles. Alternative Comics: An Emerging Literature. Jackson: UP of Mississippi, 2005.
- Highsmith 1993 Highsmith, Doug. "Developing a 'Focused'" Comic Book Collection in an Academic Library". *The Acquisitions Librarian* 4: 8 (1993), pp. 59-68.
- Hoffmann 1988 Hoffman, Frank. "Comic Books in Libraries, Archives and Media Centers". *The Serials Librarian* 16: 1 (1988), pp. 167-198.
- Huebner 2009 Huebner, Andrew J. "Secret Identity Crisis: Comic Books and the Unmasking of Cold War America". The Sixties: A Journal of History, Politics and Culture 2: 2 (2009), pp. 268-270.
- International Correspondence Schools 1968 International Correspondence Schools. ". Advertisement". In *Fantastic Four* #78. Marvel Comics, September 1968. pp. 4-4.
- Jacobs 2007 Jacobs, Dale. "More Than Words: Comics as a Means of Teaching Multiple Literacies". *English Journal* 96: 3 (2007), pp. 19-25.
- Jeet 2008 Heer, Jeet, and Kent Worcester, eds. A Comics Studies Reader. Jackson: UP of Mississippi, 2008.
- Kirby 1976 Kirby, Jack. Captain America #193. Captain America. Marvel Comics, January 1976.
- Kirby and Lee 1965? Kirby, Jack, and Stan Lee. "Manuscript for The X-Men #17". In *Jack Kirby Collector* #46. 2006. pp. 20-20.
- Kirby and Lee 1966 Lee, Stan, and Jack Kirby. The X-Men #17. The X-Men. Marvel Comics, February 1966.
- Kunzle 2007 Kunzle, David. Father of the Comic Strip: Rodolphe Töpffer. Jackson: UP of Mississippi, 2007.
- Kuskin 2008 Kuskin, William. "Graphia: The Graphic Novel and Literary Criticism". English Language Notes 46: 2 (2008).
- Lee 1941 Lee, Stan. "Captain America Foils the Traitor's Revenge". Captain America Comics #3. In Jeff Youngquist, ed., Marvel Visionaries: Stan Lee. New York: Marvel Comics, 2005. pp. 7-8.
- Lee 1961 Lee, Stan. "Fantastic Four Synopsis". In Roy Thomas and Peter Sanderson, eds., *The Marvel Vault*. Philadelphia: Running Press, 2007. pp. 67-67.
- Lee and Ditko 1962 Lee, Stan, and Steve Ditko. *Amazing Fantasy #15. Amazing Fantasy.* Atlas Magazines [Marvel Comics], September 1962.
- Lee and Ditko 1963 Lee, Stan, and Steve Ditko. Amazing Spider-Man #5. Amazing Spider-Man. Non-Pareil Publishing [Marvel Comics], September 1963.
- Lee and Kirby 1963 Lee, Stan, and Jack Kirby. *The X-Men #1. The X-Men.* Canam Publishers [Marvel Comics], September 1963.
- Lee and Kirby 1966a Lee, Stan, and Jack Kirby. Fantastic Four #51. Fantastic Four. Canam Publishers [Marvel Comics], June 1966.
- Lee and Kirby 1966b Lee, Stan, and Jack Kirby. "The Super-Adaptoid!". In *Tales of Suspense #84*. Vista Publications [Marvel Comics], December 1966. pp. 18-30.
- Lee and McLaughlin 2007 Lee, Stan, and Jeff McLaughlin. Stan Lee: Conversations. Jackson: UP of Mississippi, 2007.
- Lee and Steranko 1969 Lee, Stan, and Jim Steranko. *Captain America #111. Captain America*. Marvel Comics, March 1969.
- Lee and Trimpe 1969 Lee, Stan, and Herb Trimpe. *The Incredible Hulk #114. The Incredible Hulk.* Marvel Comics, April 1969.
- Llull and Pereira 1990 Llull, Ramon, and Michela Pereira. Raimundi Lulli Opera Latina. Turnholti: Brepols, 1990.

Matz 2004 Matz, Chris. "Collecting Comic Books for an Academic Library". Collection Building 23: 2 (2004), pp. 96-99.

McCloud 1993 McCloud, Scott. Understanding Comics: The Invisible Art. Northampton: Tundra Publications, 1993.

- McGann 2001 McGann, Jerome. Radiant Textuality: Literature After the World Wide Web. New York: Palgrave Macmillan, 2001.
- McLain 2009 McLain, Karline. India's Immortal Comic Books: Gods, Kings, and Other Heroes. Bloomington: Indiana UP, 2009.
- **Moldoff 1940** Moldoff, Sheldon. "Flash Comics #5". In Jules Feiffer, ed., *The Great Comic Book Heroes*. New York: Bonanza Books, 1965. pp. 132-141.
- Moore 1987 Moore, Alan, and Dave Gibbons. Watchmen. New York: DC Comics, 1987.
- Moretti 2000 Moretti, Franco. "Conjectures on World Literature". New Left Review 1 (2000), pp. 54-68.
- Morrison 2000 Morrison, Grant, and J.G. Jones. Marvel Boy #4. Marvel Boy. Marvel Comics, November 2000.
- Norton 2003 Norton, Bonny. "The Motivating Power of Comic Books: Insights from Archie Comic Readers". *Reading Teacher* 57: 2 (2003), pp. 140-147.
- Pratt 2009 Pratt, Henry John. "Narrative in Comics". The Journal of Aesthetics and Art Criticism 67: 1 (2009), pp. 107-117.
- Ranker 2007 Ranker, Jason. "Using Comic Books as Read-Alouds: Insights on Reading Instruction From an English as a Second Language Classroom". *Reading Teacher* 62: 4 (2007), pp. 296-305.
- **Rust 2008** Rust, Martha. "It's a Magical World: The Page in Comics and Medieval Manuscripts". *English Language Notes* 46: 2 (2008), pp. 23-38.
- Selznick 2007 Selznick, Brian. The Invention of Hugo Cabret. New York: Scholastic, 2007.
- Simon and Kirby 1965 Simon, Joe, and Jack Kirby. "Captain America Comics #1". In Jules Feiffer, ed., *The Great Comic Book Heroes*. New York: Bonanza Books, 1965. pp. 163-170.
- Sperberg-McQueen, Huitfeldt, and Renear 2000 Sperberg-McQueen, C.M., C. Huitfeldt and A. Renear. "Meaning and interpretation of markup". In *Markup Languages: Theory & Practice (Volume 2)*. pp. 215-234.
- **TEI 2010a** TEI Consortium. "sound". Appendix C: Elements. In TEI Consortium, ed., TEI P5: Guidelines for Electronic Text Encoding and Interchange. TEI Consortium, 2010. http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-sound.html.
- **TEI 2010b** TEI Consortium, ed. "Using the TEI". In *TEI P5: Guidelines for Electronic Text Encoding and Interchange*. TEI Consortium, 2010. http://www.tei-c.org/release/doc/tei-p5-doc/en/html/USE.html.
- **TEI 2010c** TEI Consortium, ed. *TEI P5: Guidelines for Electronic Text Encoding and Interchange*. TEI Consortium, 2010. http://www.tei-c.org/release/doc/tei-p5-doc/en/html/.
- Thomas 1964 Thomas, Roy. "Letter to Stan Lee and Jack Kirby". *Fantastic 4 Fan Page*. In *The Fantastic Four* #22. Marvel Comics, January 1964. pp. 30-30.
- Wanzo 2009 Wanzo, Rebecca. "Wearing Hero-Face: Black Citizens and Melancholic Patriotism in *Truth: Red, White, and Black*". *The Journal of Popular Culture* 42: 2 (2009), pp. 339-362.
- Whitson 2007 Whitson, Roger, and Donald Ault, eds. "William Blake and Visual Culture". *ImageText* 3: 2 (2007). http://www.english.ufl.edu/imagetext/archives/v3_2/.
- Wright 2001 Wright, Bradford W. Comic Book Nation : The Transformation of Youth Culture in America. Baltimore: Johns Hopkins UP, 2001.

This work is licensed under a Creative Commons Attribution-NoDerivatives 4.0 International License.