An Emerging Canon? A Preliminary Analysis of All References to Creative Works in Critical Writing Documented in the ELMCIP Electronic Literature Knowledge Base

(ELO 2013 Conference paper, August 2013 Draft)

Scott Rettberg

University of Bergen Electronic Literature Research Group

Abstract

As of July 2013, the ELMCIP Electronic Literature Knowledge Base includes documentation of more than 2,000 creative works and more than 2,000 articles of critical writing. Many of the records of critical writing include cross-references to the creative works they address. This article presents a preliminary analysis of all of the critical writing-to-creative work cross-references currently documented in the Knowledge Base in the aggregate. By developing static and interactive visualizations of this data, we might begin to see the outlines of an emerging "canon" of electronic literature.

Introduction

Every time contributors add a record to the ELMCIP Electronic Literature Knowledge Base, they have the opportunity to add references to creative works of other articles of critical writing referenced. This enables the formation of a network of critical relations, what we have described in the ELMCIP Knowledge Base project report as a "literary ecology." Using node references and attached views in the databases, these cross-references automatically display on both the record for critical writing and creative work it refers to. Over time, this develops into documentation of the critical reception of any given work documented in the Knowledge Base. While these relations are visible on the local level of the individual record, they can also be harvested from the database in filtered groups or as a whole. The nodes and the relationships between them can then visualized using software such as Gephi, to enable new ways of seeing patterns in the data and navigating through it[1]. The network visualizations presented here represent views of all of the creative works documented in the Knowledge Base that have been cited in works of critical writing. While this is still a relatively preliminary picture, one could assert that they represent an impression of the "canon" of electronic literature, if one were to presume that a literary canon is based on the critical attention paid to literary artifacts. This article

---

[1] The Bergen Electronic Literature Research Group's first forays into using the Knowledge Base as a data source for macroanalytic visualizations are documented in "Mining the Knowledge Base: Exploring Methodologies for Analysing the Field of Electronic Literature", Rettberg and Walker Rettberg, 2013.

presents some preliminary results of visualizing and reading the relationships between critical writing and the works that they reference, as connections that together define a field.

Literary canonization is always a problematic subject, and any effort to construct a particular canon will pose challenges to its selectors. Astrid Ensslin's *Canonizing Hypertext* (Ensslin 2007) is the scholarly work in the field of electronic literature that has most assiduously addressed these challenges. Citing Winko, Ensslin notes that canons "comprise texts which are considered culturally valuable by a certain group or society and therefore 'worthy' of being handed on to posterity." (48) Ensslin asserts that these canons have important psychological and socializing effects on members of the group familiar with the works. She further explains that canonization is a dynamic process, and that in multilayer societies "alternative canons frequently undermine imposed canons. After thoroughly considering conceptual and historical frameworks of canonization, she then asserts and describes her own "alternative canon" to the mainstream: one for literary hypertext. Ensslin then proposes a set of criteria and actually produces a list of twenty-three works that she considers to be worthy of inclusion in the hypertext canon. She asserts that her selections are based on a "rule canon" that considers each literary artifact according to axiological criteria, addressing aspects of Production (Innovativeness), Object (Thematic depth), Form (Aesthetic overstructuring and Semiotic interplay) and Reception (Criticism, Anthologization/degree of canonization and Motivation/effect on the reader).

Hypertext and the Question of Canonicity		65

Table 2 'Rule canon'

Production	Innovativeness	• deviation from literary/hypertextual traditions • interrelation between technology and subject matter
Object	Thematic depth	• topicality • thematic message • self-reflexivity • metatheoretical concept • intertextuality
Form	Aesthetic overstructuring (microstructurally/ macrostructurally)	• rhetorical devices • linking patterns • navigation • hypertext structure
	Semiotic interplay	• transmedialization • implemented intertextuality
Reception	Criticism	• critical acclaim • awards
	Anthologization/degree of canonization	• readers (print/CD-ROM) • university reading lists • curricular presence
	Motivation/effect on reader	• suspense • surprise • playfulness • 'intellectual exercise'

Fig 1. Ensslin's 'Rule canon' criteria for canonization of hypertext (65).

Ensslin's criteria appear to be well-thought out but quite purposefully leave a great deal of room for interpretation, and in the case of Ensslin's study, interpretation performed by a solitary critic working in relative isolation. There is great value in critical interpretation, but anyone who sets out alone to define a given canon for an entire field of practice is taking a bold, or even arrogant, critical position. The critical reception of a work might be the one area in which we could hope for a clearly empirical measure of canonicity. Yet Ensslin asserts a particularly privileged position with regard to reception when she writes that:

> The research situation with most hypertexts is such that reviews and academic papers
> are written by hypertext supporters. Therefore, criticism tends to be rather opinionated
> and to emphasize the academically interesting sides of a hypertext rather than its

> cumbersome attributes. Evaluations on aesthetic quality from my own point of view are therefore added to reinforce, supplement or indeed oppose other critic's appraisals. (69)

In other words, Ensslin is saying that in constructing a hypertext canon, her own judgment will reign supreme, and the judgment of other critics in the field are hers to pick and choose from, to amplify or discard at will: her criteria, her judgment, her prerogatives, her canon. While I appreciate Ensslin's careful formulation of her criteria for canonization, one must nonetheless recognize that it functions better as a framework for the personal scholarly *production* of a canon than it does as a framework for understanding a canon that actually exists in a more widespread frame of critical practice. Ensslin hedges this a bit, asserting that that her "hypertext canon is only single-handed with regard to its selector anyhow. The actual selection process operates by means of (electronically) networked reading and scholarly discourse … the text catalogue reflects an individual opinion derived from many other opinions." But this still begs the question: is the canon *in effect* the opinion of any individual (human) scholar, or is it the aggregate of all those judgments, operating independently as nodes in vast scholarly network? In other words how could we scientifically assess not what *should be in the canon* according to one critic (however well-informed he or she might be) but instead what *actually is in the canon* by some more objective measure? That is to say that while I, or Astrid Ensslin, or Roberto Simanowski, or Philippe Bootz, or N. Katherine Hayles, or Nick Montfort, or any number of other critics of electronic literature no doubt have our own subjective shortlists of digital literary artifacts worthy of preservation, greater circulation, teaching and so forth and each have our own values and selection criteria, trying to understand the critical work around electronic literature as an aggregate may be worthwhile. What understanding of a field can looking at many critics' references put together produce? The ELMCIP Electronic Literature Knowledge Base can offer us some new tools to consider this question more empirically than we have before been able to.

There is a larger question regarding canonicity: is the concept of a canon actually useful to this field of practice? If traditional literary canons are intended to select out, highlight and mark for historical preservation certain printed literary works, do we even think that is necessary in a field in which the material processes of production and preservation are markedly different from those of print literary culture? Many of the twenty-three works that Astrid Ensslin identified on her hypertext canon are already very difficult to access. As Alexandra Saemmer notes "Digital literature resists any claim for physical permanence because of the lability of the device" (Saemmer 2012). We can be virtually certain that some important works will not survive

because of the contingencies of the media in which they are produced, because of their dependence on software and hardware platforms that are subject to cycles of obsolescence. Yet conversely, these very contingencies might drive a process of canonization: because a good deal of effort is presently required to document and preserve works of electronic literature and because the materials involve include an incredibly diverse range of technical formats, it may be worthwhile to speed the plow and attempt to select works for preservation before the opportunity to do so is lost. Saemmer notes that Joergen and Schaeffer "wonder if digital works generate a brand new literary value and if this value could possibly challenge the traditional aesthetic claims to perfection, consistency and stability" and asserts that they in fact do, calling for careful and continuous re-assessment of criteria for "literariness" in anthology projects and for redefining or dropping the notion of canonicity altogether. Saemmer focused on the importance of electronic literature anthologies and projects such as the enhanced fiches in NT2's *Répertoire des Arts et Littératures Hypermédiatiques*, as these result to some extent in the preservation, greater dissemination, and enhanced documentation of the works that are selected for inclusion.

I agree that we should take any proposition of a canon for electronic literature with a large chunk of salt. Indeed, though I was one of the editors of the first volume of the *Electronic Literature Collection* (2006), and agree that including a work in an anthology lends it a kind of imprimatur and importantly increases the likelihood that it will be widely circulated and preserved, I have consistently argued that we should not view the publication of works in any given anthology as a defining a canon, certainly not one that is intended to be understood as fixed for posterity. In my view, these publications are just one type of action in a complex unfolding process of collective field building that spans decades and cannot be reduced to singular events of publication. If certain works are selected for special attention as exemplary or representative at a given moment, the field is not yet at a point where it would benefit from the type of closure that the concept of a canon may imply.

In comparison to developing a subjective list of favored works or considering the selection of works for the two *Electronic Literature Collections* or for the NT2 enhanced fiches, the relative reception of works of electronic literature might be better measured by observing how they circulate in critical discourse. The cross-references in the ELMCIP Knowledge Base can offer an impression, albeit an imperfect one, of how this circulation is taking place.

Developing a macro-level understanding of how critical discourse has focused on particular works is useful, but not necessarily as a measure of what should be set aside for posterity. This distant reading of a critical field might better enable us, for instance, to identify genres of creative and theoretical practice, to see how critical attention and focus has changed and diversified over time, to identify areas in which critical attention is lacking (indeed asking what works have not garnered critical attention and why may be fruitful) and to understand in a more objective longitudinal way the development and history of the field in which we practice.

Methodology

As Ensslin's schema demonstrates, most considerations of the relative canonicity of literary works would take into account more factors than simply counting the number of references to the work. With this rather brutish approach, one cannot for example distinguish the difference between positive and negative criticism of any given work. We are essentially counting links from one type of record in a database to another type of record. All that is being measured is how many items of critical writing documented in the database reference any given creative work in the database.

It is important to emphasize that the sample of the field that we have at our disposal in the ELMCIP Knowledge Base is nowhere near complete. While it is possible for every record of critical writing about electronic literature that is added to the Knowledge Base to include cross-references, not all of them do: while there were about 2,000 critical writing records in the Knowledge Base on July 6, 2013, when the data for this study was downloaded, 601 of those records included references to creative works: not all of them should have such references, as some of the records are books or articles that are referenced by other articles of critical writing and do not themselves reference electronic literature, and others, such as records based on conference abstracts, are difficult to extensively document. But I would expect that when the critical records were complete, at least 60% or so of them will reference creative works: right now only about 30% do. Many of the records in the Knowledge Base are "stubs" or incomplete records, and more still include incomplete references. Contributors add new records and update existing records on a near-daily basis, so the statistics and visualizations presented here will show a changing picture of the field each time they are derived.

There is some built-in bias to this sample: perhaps 50-60 people have contributed records to the database. The vast majority of the records have been contributed by members and guest

researchers of the Electronic Literature Research Group at the University of Bergen, and the works and critical writing we are most familiar with, including those we have published ourselves, are doubtlessly over-represented. Nevertheless we have tried to be diligent and systematic about cross-referencing works from critical writing. The sources of the critical writing records include the majority of the most significant monographs and multi-author collections of essays in the field, more than thirty dissertations about electronic literature, articles published in major journals such as *Dichtung Digital*, *ebr*, *Cybertext Yearbook*, and *The Iowa Review Web*, conference papers presented at the ELO conferences, ACM Hypertext conferences, E-Poetry Festivals, and others, as well as other forms of scholarly discourse such as reviews, interviews, video interviews, and other sources. None of this information gathering is complete, nor could it ever be within a contemporary field that is more active today than it has ever been before.

It is both a strength and flaw of the ELMCIP Knowledge Base that humans contribute all of this information, rather than automatically harvesting it, as is the case in some other types of reference databases. While a semi-automated solution may be possible in the future, many of the referenced articles include only references to criticism in their works cited list, and none to the creative works discussed. This is a bad practice that I hope critics will discontinue. Creative works should certainly be included in the bibliography of critical writing that addresses them. In any case, many of these references would have been missed if bibliographies were scraped automatically rather than handled by human contributors.

Incompleteness and biases asides, a strong cross-section of critical work in the field is represented, and the Knowledge Base is likely the most extensive resource documenting the connections between creative and critical practice in electronic literature currently available. The dataset I am working with includes 601 different items of critical writing referencing 1,201 creative works. 2,694 references in total are documented.

Any of the many fields in the Knowledge Base can be outputted from the Drupal installation in bulk raw data format according to specifically configured views. In this case, I outputted the Node IDs, titles, year of publication, and authors of all items of critical writing that included references to them, Node IDs, tiles, year of publication, and authors, and the "edges" (reference links) between the nodes. This dataset is the basis for all of the visualizations and analysis that follows.

Visualizations and Analysis

I exported the data from the Knowledge Base in CSV format, imported it into Gephi network visualization software for analysis, and produced three visualizations. The first includes all of the nodes and edges in the set. A Force Atlas 2 layout algorithm was applied, which can be useful for identifying network communities. In this case however, because there are so many nodes and so many common references across different types of criticism, the network community information is of little semantic value. The most visible aspect here is that particular sources of critical references are hubs, such as Leonardo Flores' I ♥ E-Poetry project (the node with many edges emanating from it appearing in turquoise on the lower left of the image), or the *ELMCIP Report on Electronic Literature Publishing and Distribution in Europe* (in purple to the right above *afternoon, a story*), each of which included many references to creative works. The clustering here isn't very informative. More useful here is the sizing of creative work nodes and labels, which indicates their in-degree—the number of references to them from critical writing.
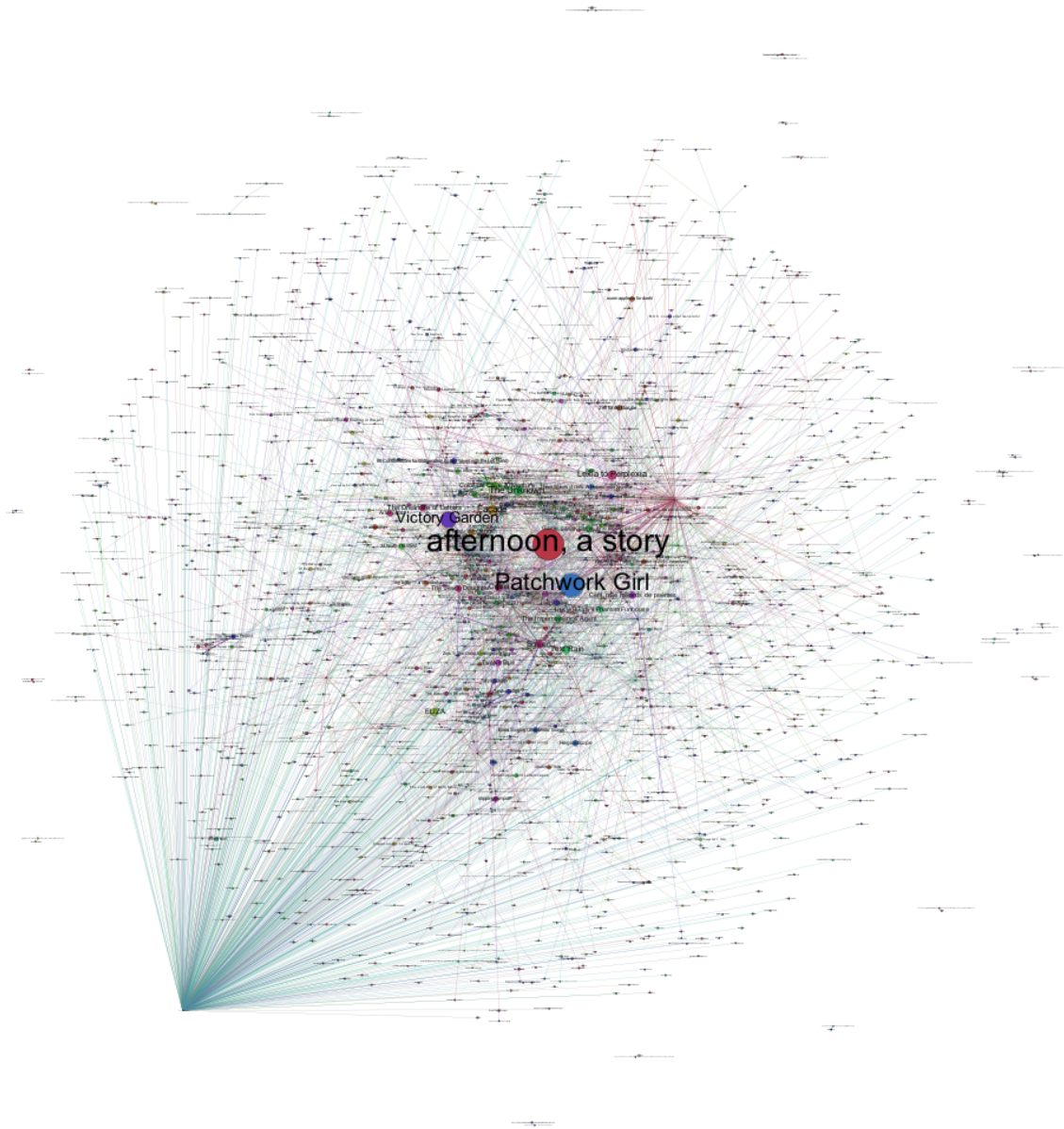
Fig. 2 All citations of creative works in critical writing documented in the ELMCIP Knowledge Base. See http://retts.net/viz/elmcip_all_cite_rev.pdf for detailed zoomable PDF of this image and http://retts.net/viz/elmcip_all_crit_citations_of_work/ for a searchable, interactive web version.

As is often the case in visualization-based research, after spending some time with this first iteration of the visualization, I concluded that it does not in itself offer a great deal of insight, except to establish that a few works are clearly the most-often cited. The sigma.js Gephi plugin allows for exporting the network graph to an interactive web version. This version is somewhat more useful, as the graph itself can both be easily zoomed for detail and becomes itself a

searchable database. The user can also select any given creative work node and see at a glance all of the critical writing references to it as in the below example of *Patchwork Girl*.
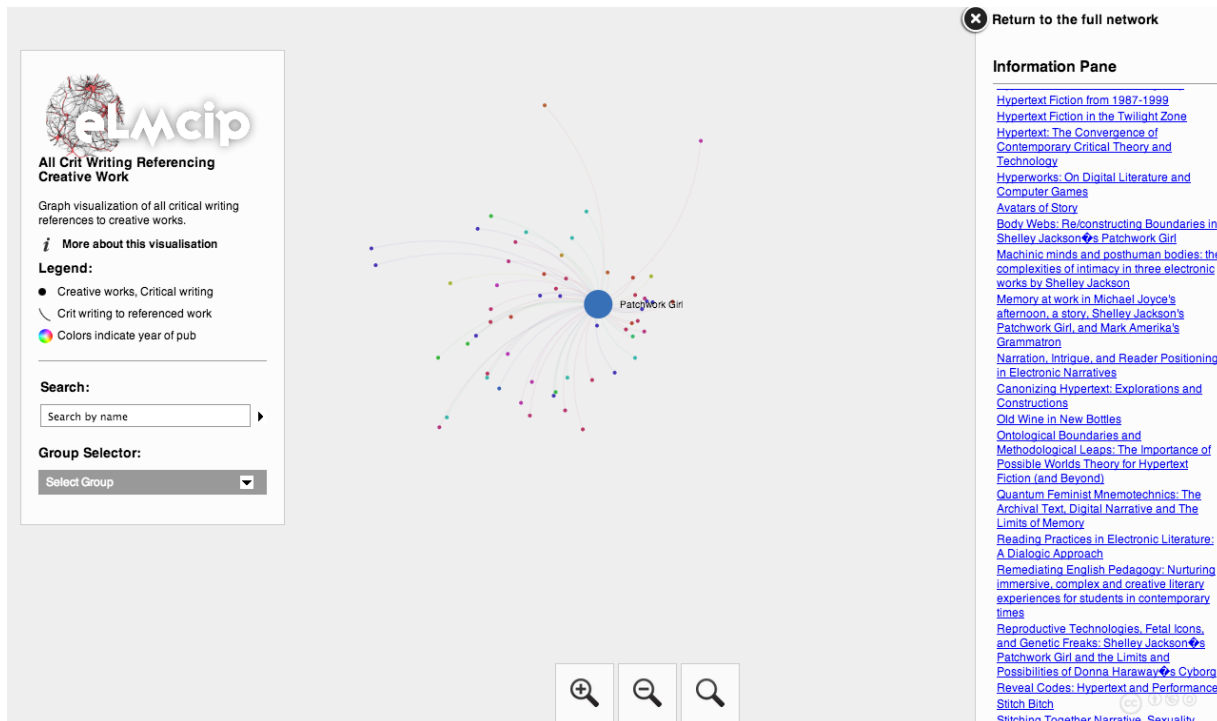

Fig. 3 Clicking on a given node of the web version of the visualization localizes the selected work and shows all critical writing nodes that reference it.

After producing the interactive web version of the visualization, I elected to pursue a different avenue of inquiry with the same pool of information. Since I pulled in year of publication as one of the node properties, temporal mapping is possible as well. While keeping creative work nodes sized by node, I colored and sorted them by year of publication.
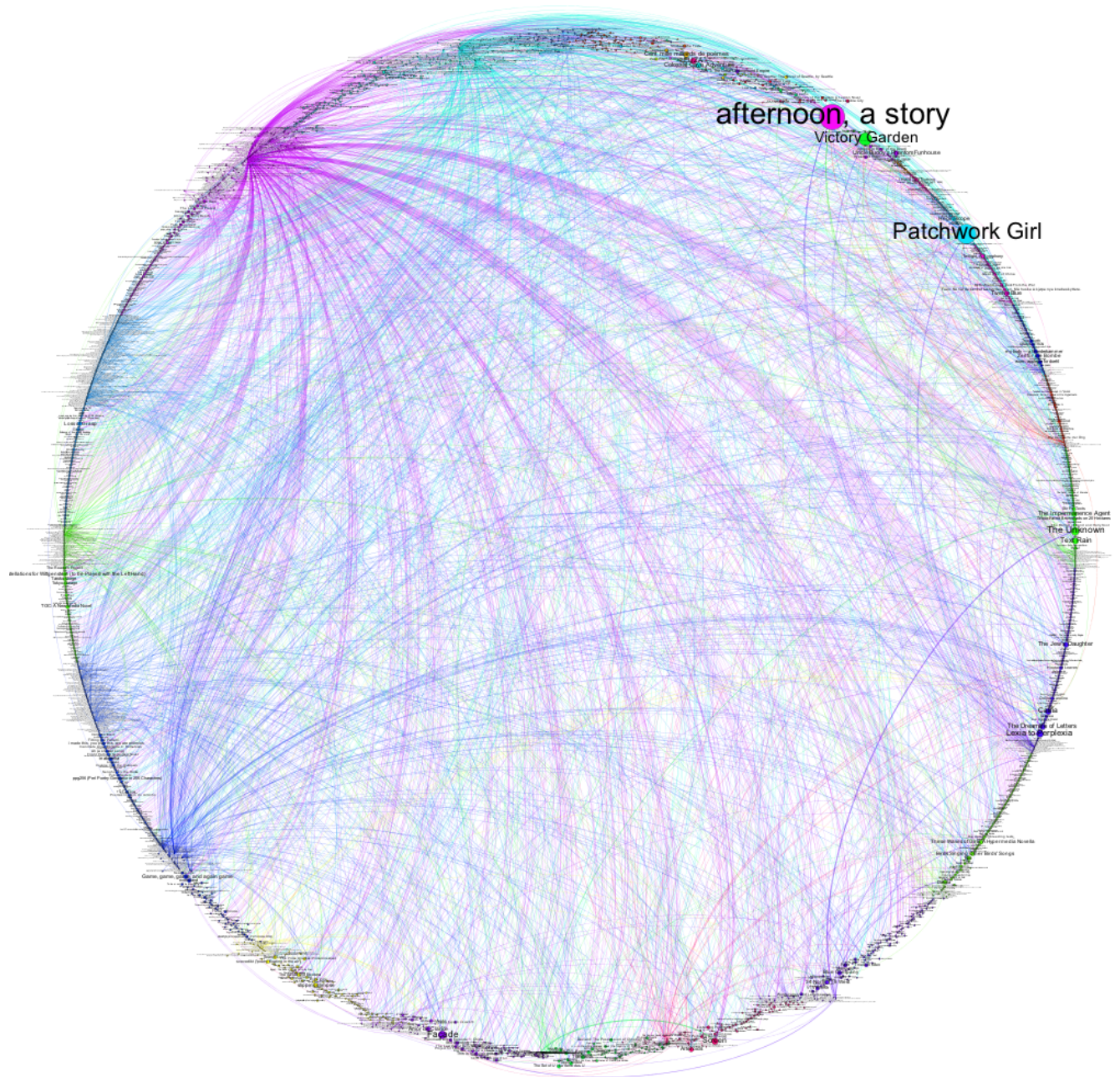
Fig. 4 All citations of creative works by critical writing in the ELMCIP Knowledge Base. Nodes and labels of creative works are sized to indicate number of references to them. Node colors and arrangement on the circle indicate year of publication. Earlier dates are to the right of the 12:00 position on the circle and proceed clockwise. A poster-sized image of this visualization is available for download at http://retts.net/viz/circ_all_citations_xl_13072013.pdf.

A large high-resolution image of this graph is useful in communicating a great deal of

information in one space, both about the content of the database and about the most-referenced

works. The temporal distribution in particular reveals the concentration of records in the database around recent years.
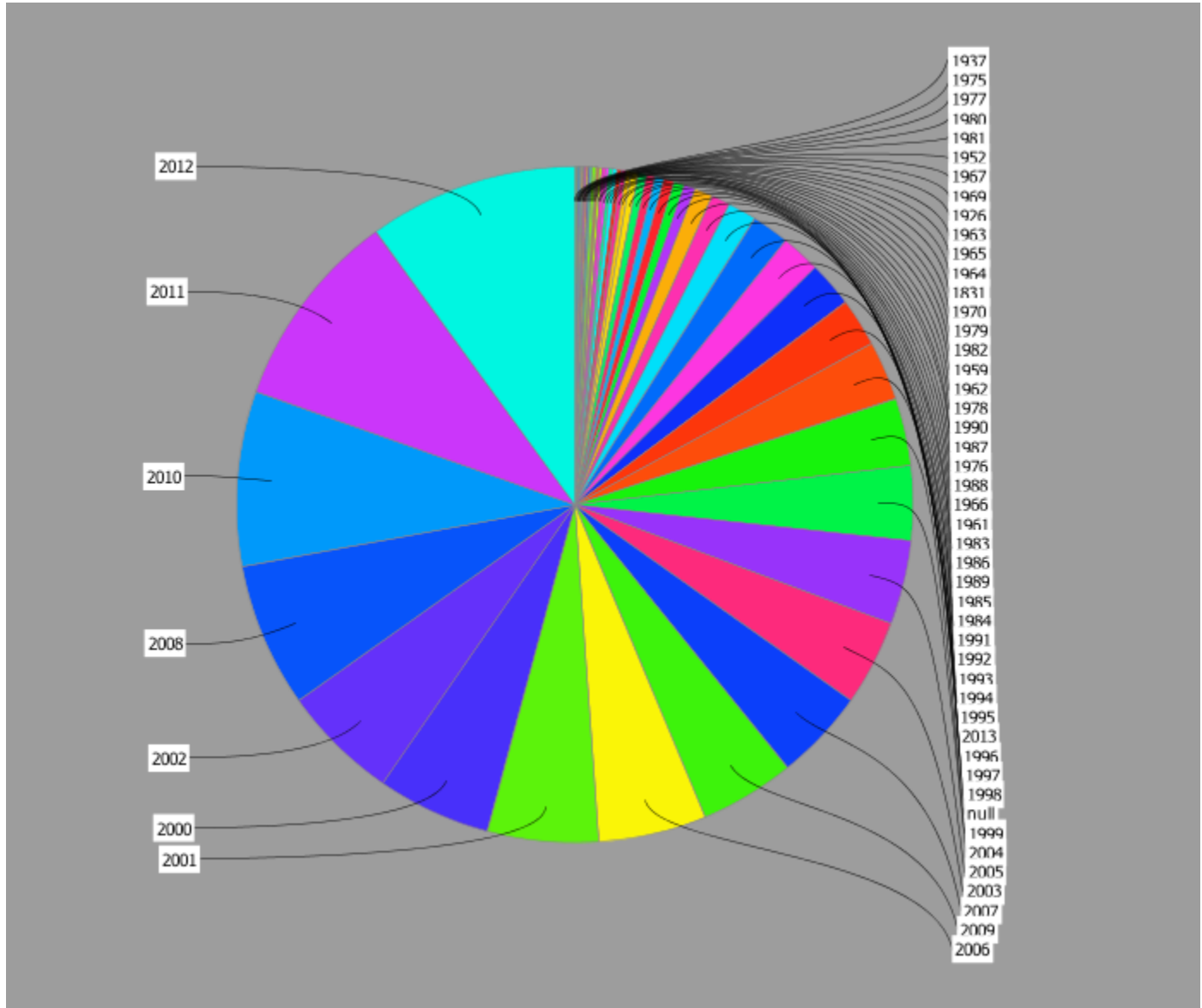

Fig. 5 Pie chart showing distribution of years of publication of the works and critical writing records in the dataset.

Not surprisingly, the years of publication of works and critical writing with the most records in the Knowledge Base are those in which the HERA-funded ELMCIP joint research project took place, 2010-2013, including more than 25% of the records. About 65% of the records are for works or critical writing published since 2000. While we can say for certain that this represents the distribution of records in the database, it also indicates a general increase in critical and creative

production within the field in recent years. Better mapping of historical activity in the field will likely shift this balance somewhat, but the field has had a more clearly organized structure in the post-millennial years. A number of academic organizations have formed during this period and held regular conferences where academics and writers gather to present research and new creative works.

Having noted this concentration of creative and critical activity in the post-2000 years, it is nevertheless the case that in the majority of the most-cited works in the field were published in earlier years, with the "holy trinity" of three works published by Eastgate Systems: *afternoon, a story* by Michael Joyce (1990), *Patchwork Girl* by Shelley Jackson (1995), and *Victory Garden* by Stuart Moulthrop (1991) the most-cited works by far in the dataset. This is likely due both to the fact that they are high-quality works of enduring interest to scholars and to the fact that much of the earlier criticism of hypertext fiction tended to center around a smaller overall set of works. If we look at a limited set of the years of publication of the most-cited works in the database—those that have been cited by eight or more items of critical writing—we can see even though the dataset is dominated by recent works and criticism, in a subset of forty-six most-cited works, those published in 2000 and earlier are still dominant in terms of the number of critical references.
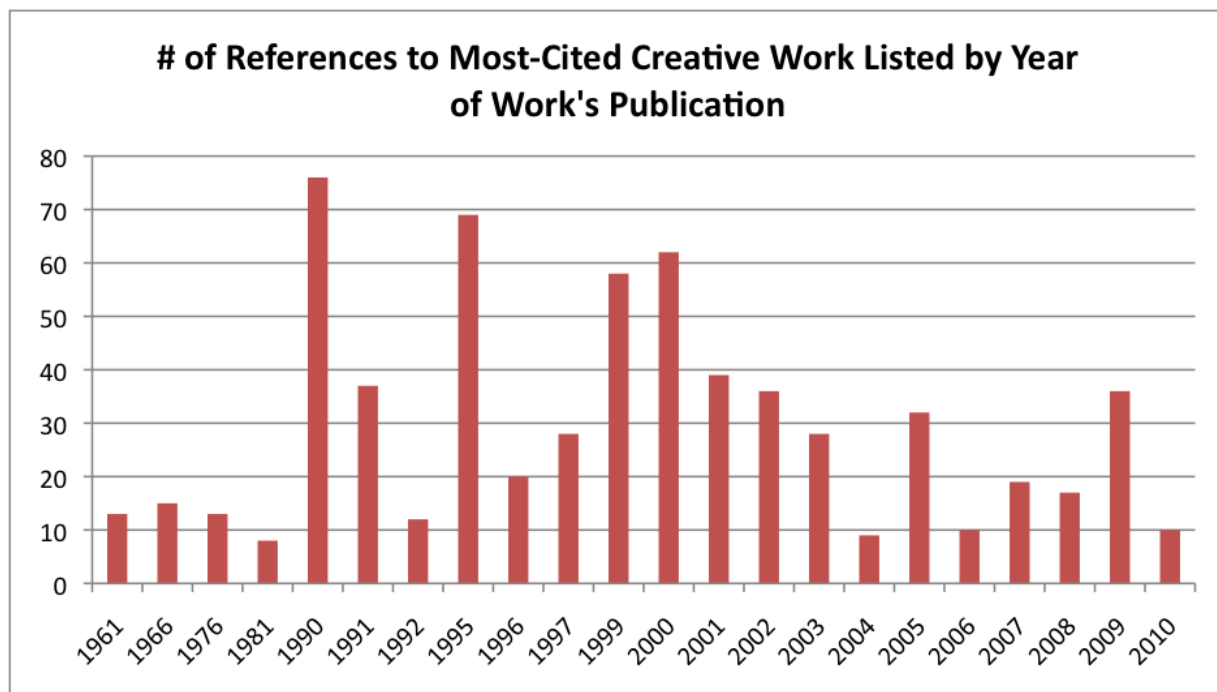

Fig. 6 Bar chart graphic representing the number of references to creative works by the work's year of publication in subset of 46 most-cited creative works (those with 8 or more critical writing references).

The diversity of referenced creative works has increased with each passing year towards the present day and fewer individual works stand out as "the defining work" of any particular period in more recent years. But a set of certain touchstone works published prior to 2001 seem to be well-established due to the precedents of prior criticism: an article on hypertext is statistically likely to include some mention of *afternoon, a story*; an article on chatbots will likely reference *ELIZA*; an article on interactive fiction will likely reference *Colossal Cave Adventure* or *Zork*.

An interactive web version of the temporally organized circular graph of the dataset proves useful as a more accessible research tool than the static visualization and has potential as a research tool in its own right, as it allows the user to see the references to any given work within a temporal span.
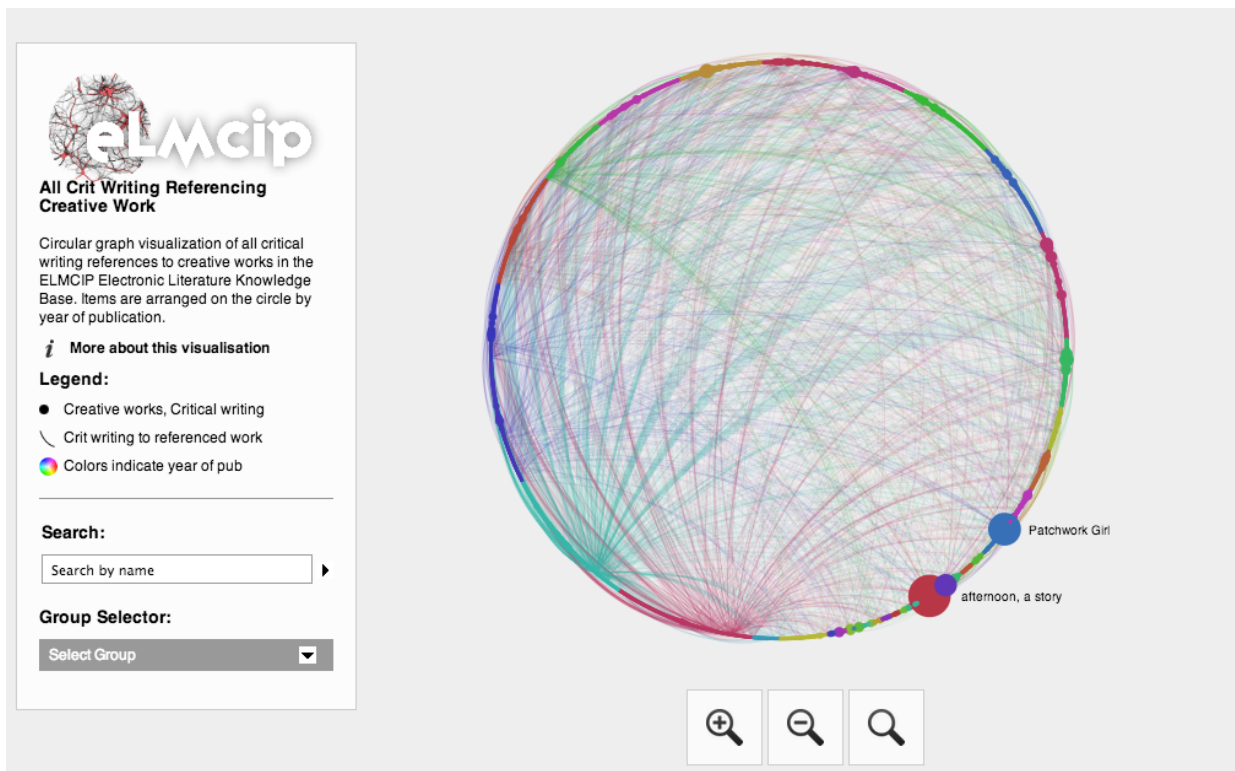

Fig. 7 Interactive web version of the circular visualization of all citations of creative works arranged by year: http://retts.net/viz/elmcip_all_crit_citations_of_work_circ/. Note that the current version of sigma.js Gephi export plugin has a bug that inverts the y axis, so in this visualization the earliest years start at the 6:00 position and increase in a counter-clockwise direction.

By selecting any given work, we can see the references to it, arranged temporally by year of publication. Though it is beyond the scope of the current study to do so in great detail, it would be worthwhile to analyze in more detail how particular works are differently received over time,

as doing so might reveal some interesting patterns in the critical practices of the field. A few examples:
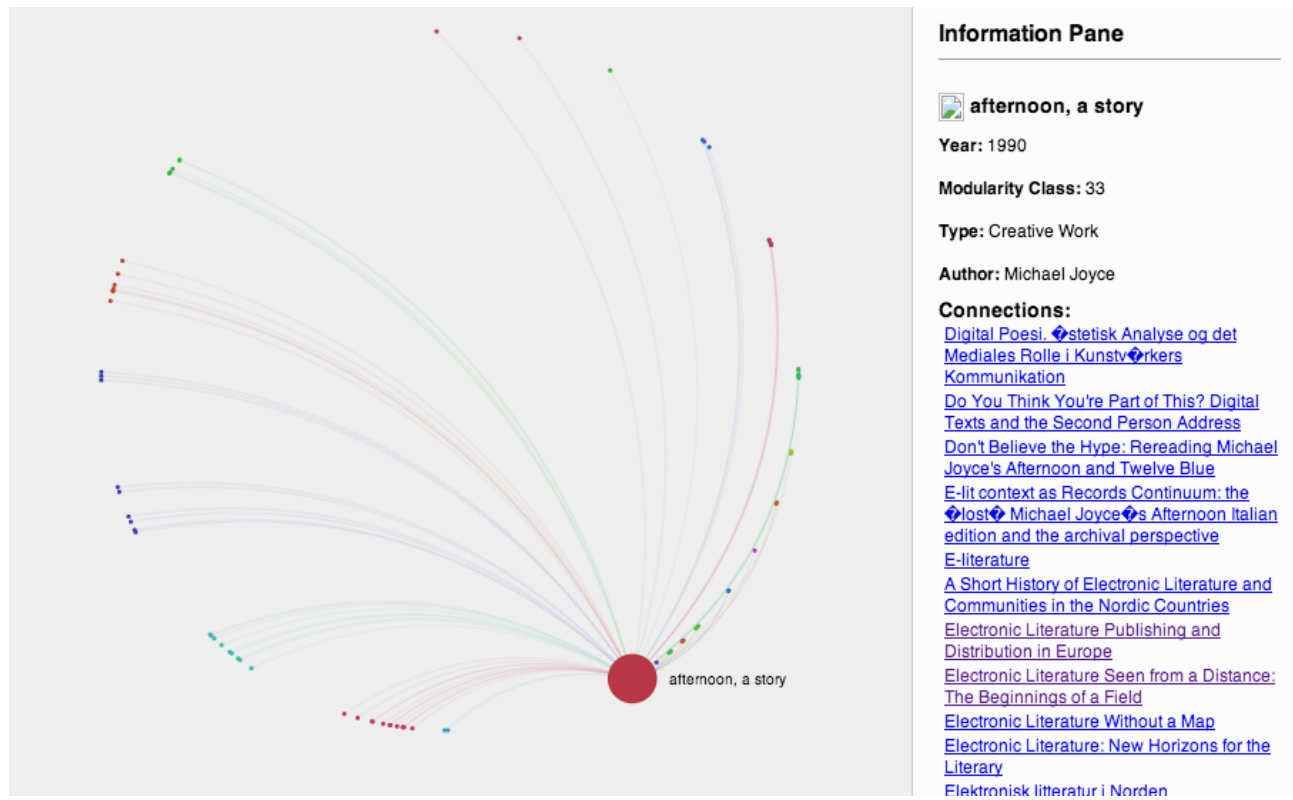

Fig. 8 Temporal view of references to *afternoon, a story* from the web visualization.
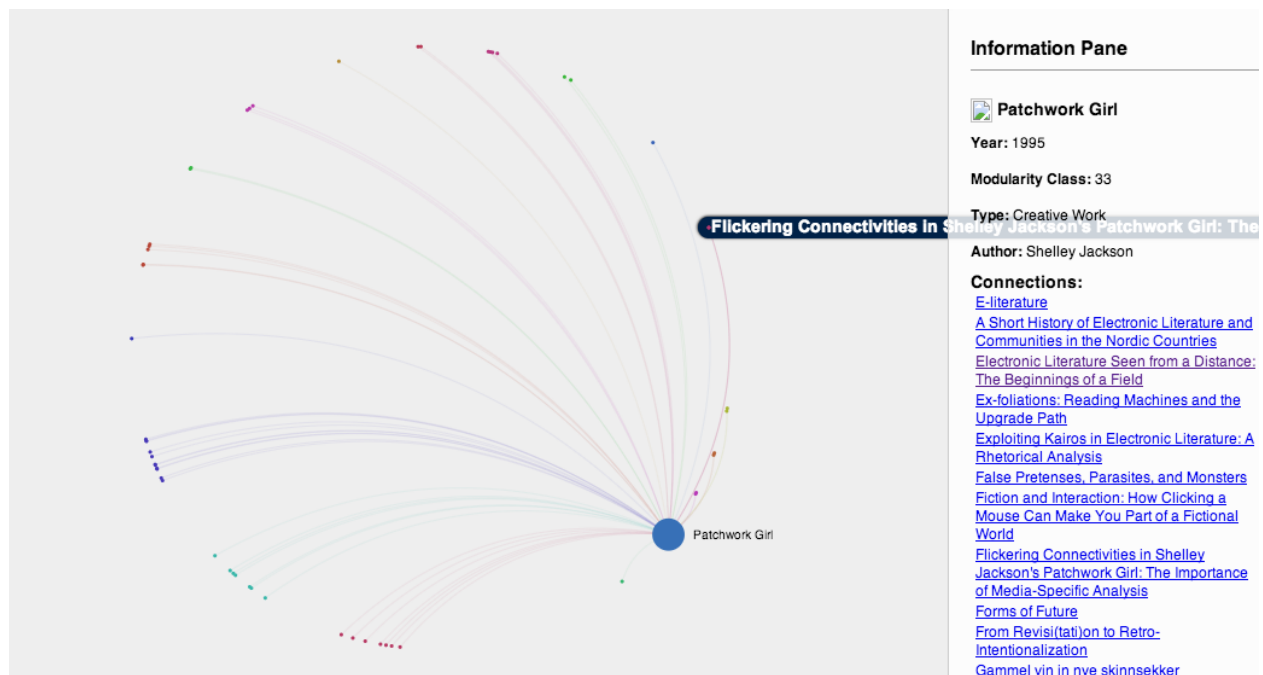

Fig. 9 Temporal view of references to *Patchwork Girl* from the web visualization.
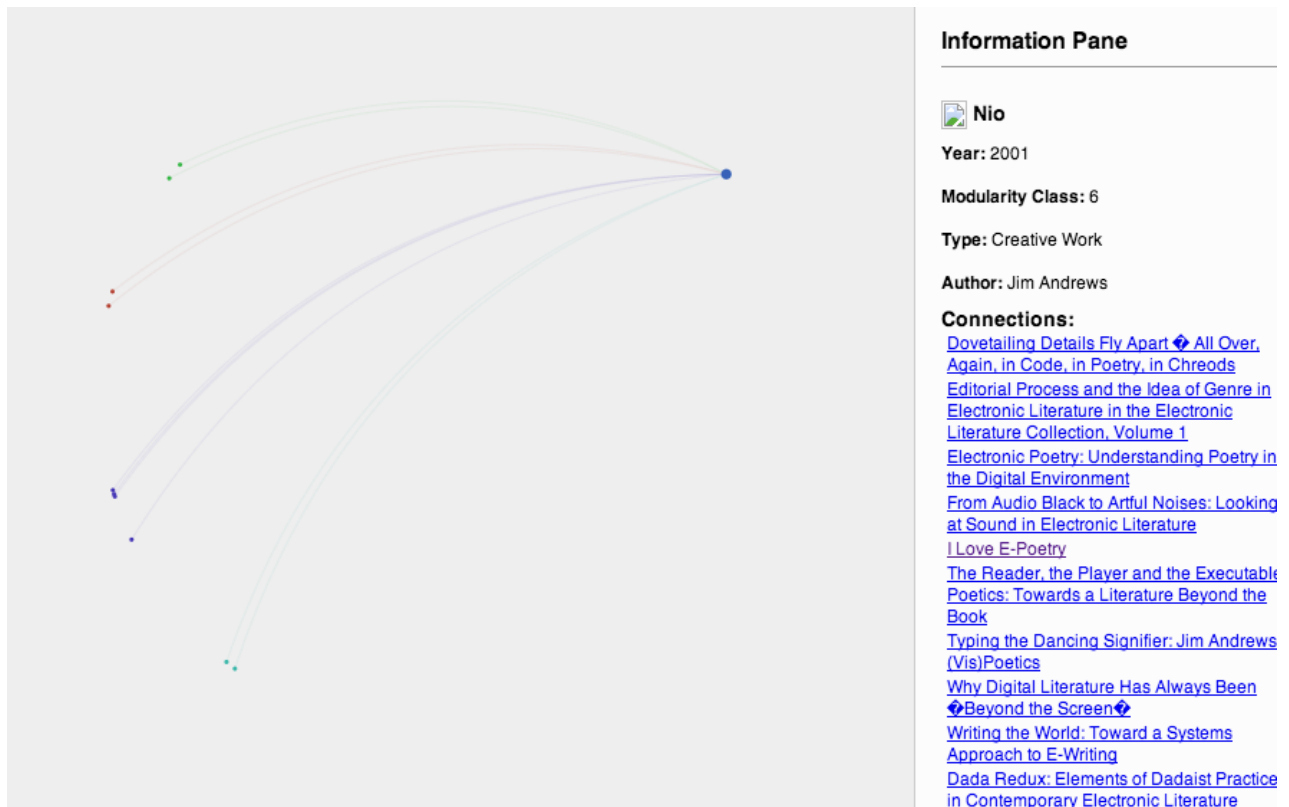
Information Pane

Nio

Year: 2001

Modularity Class: 6

Type: Creative Work

Author: Jim Andrews

Connections:
Dovetailing Details Fly Apart � All Over, Again, in Code, in Poetry, in Chreods
Editorial Process and the Idea of Genre in Electronic Literature in the Electronic Literature Collection, Volume 1
Electronic Poetry: Understanding Poetry in the Digital Environment
From Audio Black to Artful Noises: Looking at Sound in Electronic Literature
I Love E-Poetry
The Reader, the Player and the Executable Poetics: Towards a Literature Beyond the Book
Typing the Dancing Signifier: Jim Andrews (Vis)Poetics
Why Digital Literature Has Always Been �Beyond the Screen�
Writing the World: Toward a Systems Approach to E-Writing
Dada Redux: Elements of Dadaist Practice in Contemporary Electronic Literature

Fig. 10 Temporal view of references to *Nio* from the web visualization.

❌ Return to the full network

Information Pane

Fa�ade

Year: 2005

Modularity Class: 38

Type: Creative Work

Author: Michael Mateas, Andrew Stern

Connections:
Generating Narrative Variation in Interactive Fiction
Avatars of Story
Behind Fa�ade: An Interview with Andrew Stern and Michael Mateas
Beyond the Screen: Transformations of Literary Structures, Interfaces and Genres
Interactive Drama, Art, and Artificial Intelligence
Konservativer Avantgardismus oder: Die Rettung der Kunst im Zeichen der Technik
Looking Behind the Fa�ade: Playing and Performing an Interactive Drama
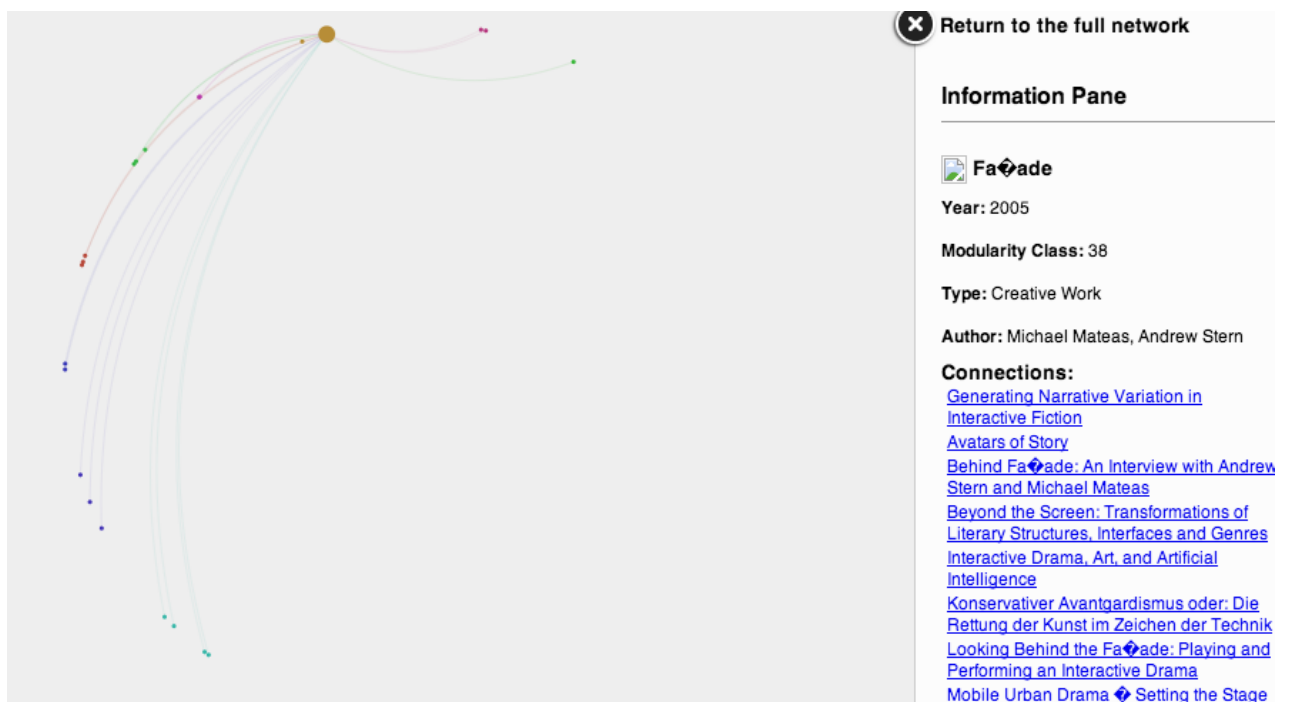Mobile Urban Drama � Setting the Stage

Fig. 11 Temporal view of references to *Façade* from the web visualization.

We can identify slightly different patterns of reception in each of these four images. Joyce's *afternoon, a story* received a great deal of sustained critical attention in the years immediately following its publication and has garnered fairly consistent critical attention ever since. Jackson's *Patchwork Girl* appears to have been written about a bit less in the years immediately after its 1995 publication, and more in subsequent years as it has been increasingly regarded as one of the classics of its genre. In a case such as this, a more detailed study might want to examine the individual critical references to see if there was a "breakout" moment of critical reception— perhaps N. Katherine Hayles' 2000 essay "Flickering Connectivities in Shelley Jackson's *Patchwork Girl*: The Importance of Media-Specific Analysis"—that inspired other critics to renew their critical interest in the work. The case of Jim Andrew's *Nio* is also interesting to consider. All of the critical writing references to it recorded in the ELMCIP Knowledge Base took place after 2006, though it was initially published in 2001. In 2006 it was published in *Electronic Literature Collection*, Volume 1. *Nio* is a work of visual poetry, in which the user acts as a kind of VJ, remixing lettristic animations and musical loops. Perhaps because it is one of the boundary cases of electronic literature, it may have taken the imprimatur of anthology publication to bring it into circulation in the scholarly discourse of electronic literature. Michael Mateas's and Andrew Stern's *Façade* is another interesting case to consider. In this case, several scholarly publications related to the work took place for several years *before* its 2005 publication. This can be explained by the fact that before the work was complete, Mateas and Stern both exhibited the work-in-progress at conferences and festivals and wrote and published articles related to its production in journals. In the years since, the work has become a familiar reference not only in publications related to electronic literature, but also game studies and artificial intelligence. The long development and incremental release process might offer some lessons for authors developing similarly ambitious e-lit projects of scale: the pump was already well primed for scholarly reception of the work before it was officially released.

Though these are just a few initial examples of how taking the temporal scale of scholarly reception of individual creative works into account might be useful, I hope they reveal how visualizations of this kind can serve not only as illustrations of research, but as research tools in their own right. Each successive visualization of the data opens up, rather than forecloses, new research questions.

After developing a set of visualizations that provided a view of all of the citations of creative works in the database, I decided to look at a smaller subset of the most-cited works to see if this

reduction would reveal any patterns or network communities. The following visualization filters the set of works to those with at least four critical writing citations in the Knowledge Base. A Force Atlas 2 layout was applied to this subset. Nodes and labels are sized by number of references, and nodes and edges are colored by modularity.
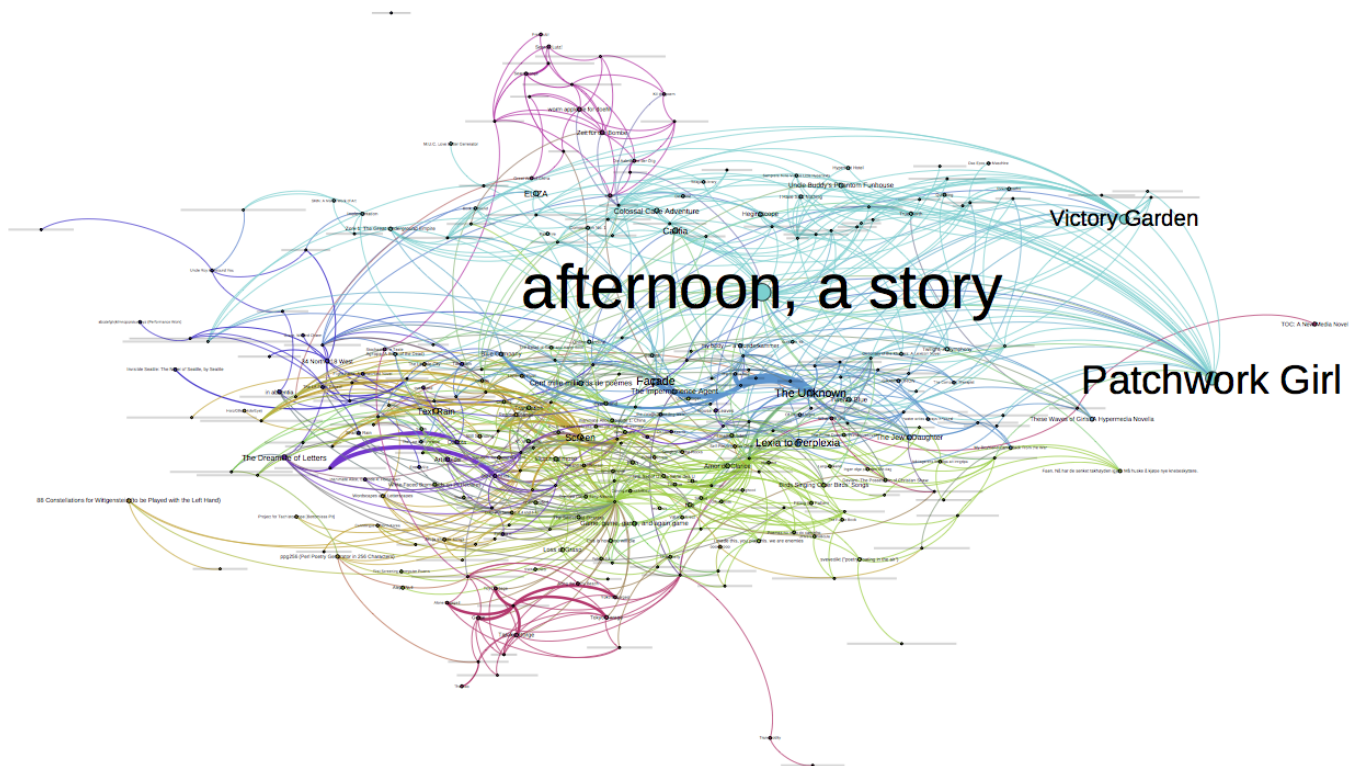

Fig. 12 Visualization of creative works in the Knowledge Base with 4+ critical writing citations. Nodes and labels are sized by number of references. Nodes and edges are colored by modularity. A zoomable PDF of this image is available at http://retts.net/viz/elmcip_works_freq_cited/works_freq_cited.pdf

The modularity algorithm attempts to find network communities and to show how the network can be compartmentalized into sub-networks or clusters. Looking at a smaller, more densely connected set of works rather than at a larger set that is more diversely distributed can offer more insight into possible groupings or genres of work. So in this case, items in any given cluster will tend to have been often cited in common in critical writing. Jill Walker Rettberg's "An Analysis of Dissertations About Electronic Literature" offers more insight into network analysis, examining a more tightly clustered subset of the critical writing in the Knowledge Base, the Ph.D. dissertations documented there.

In comparison to the visualizations of the whole dataset of referenced works, more communities of practice are identifiable here by modularity.
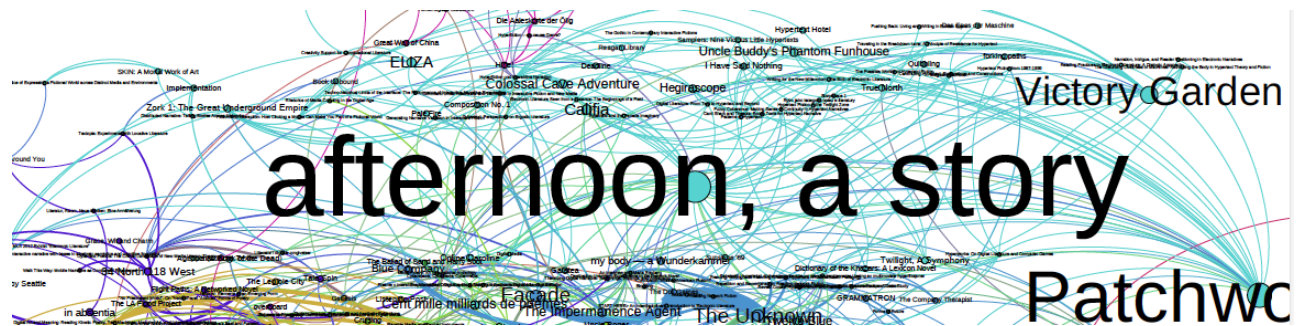

Fig. 13 Detail of visualization -- classics cluster in turquoise.

One of the identified modules is clearly a "classics of electronic literature" cluster, including Eastgate hypertext fictions from *afternoon* to *Califia* as well as early classics of other genres, for instance the chatbot *ELIZA*, and text adventures *Colossal Cave Adventure* and *Zork*.

Fig .14 Detail of visualization -- late 90s/early 2000s web narratives cluster in dark blue.

Another cluster identified is a set of web hypertexts and other narratives that appeared on mainly in the millennial period from the late 1990s until early 2000s. Hypertext narratives such as *The Jew's Daughter*, *The Unknown*, and *Lexia to Perplexia* appear alongside other narratives distributed on the web, such as the interactive drama *Façade* and the email novel *Blue Company*.

These two clusters may indicate that there is a good deal of clustering on the basis of temporality—works written about during the same period tend to be mentioned in the same contexts, though there is also a strong pull towards modularity by genre or type. The language

the work was written in also tends to have an effect on where and how it received, as shown by the "German cluster":
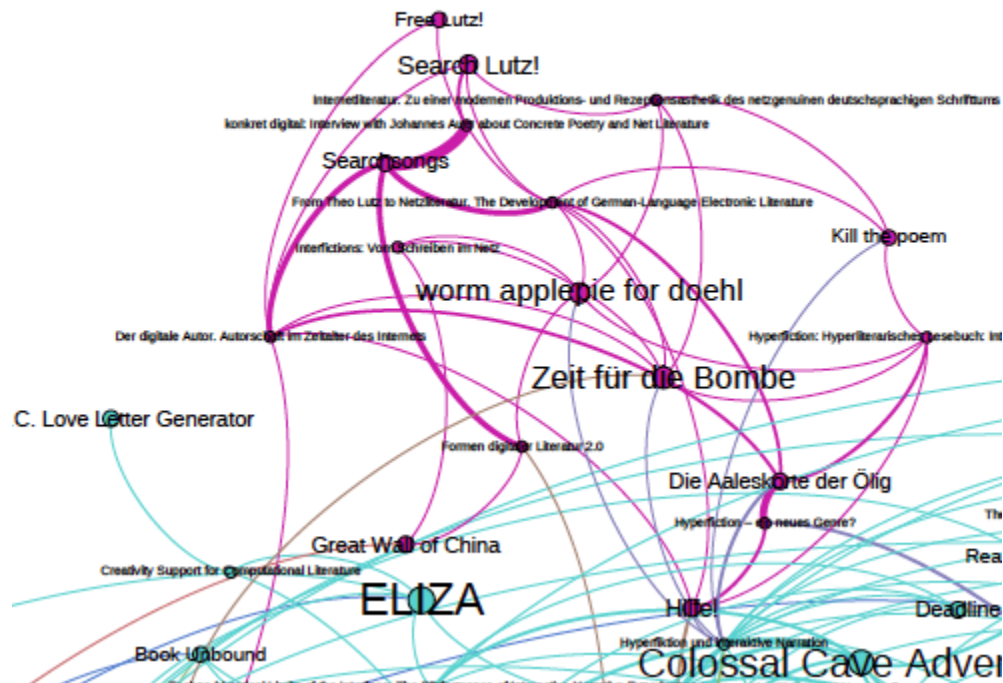

Fig .15 Detail of visualization -- late German cluster in dark pink.

Again, the visualizations here reveal the limitations as well as the strengths of the current dataset. On the basis of the existing body of creative work and criticism of e-lit in French, one would likely expect there to be a similar cluster of French works on the list of most-referenced works. The Knowledge Base is still relatively lacking in French contributors, so while there is a growing body of records for French work, we need the participation of more native-speaker French scholars in order flesh out the picture of the work and scholarship of the French e-lit scene.

The relationship of software platforms to works may also form the basis of clusters. One of the identifiable clusters is a set of works developed in Flash or similar software that emphasizes motion graphics, kinetic texts, or innovative interface and interaction.
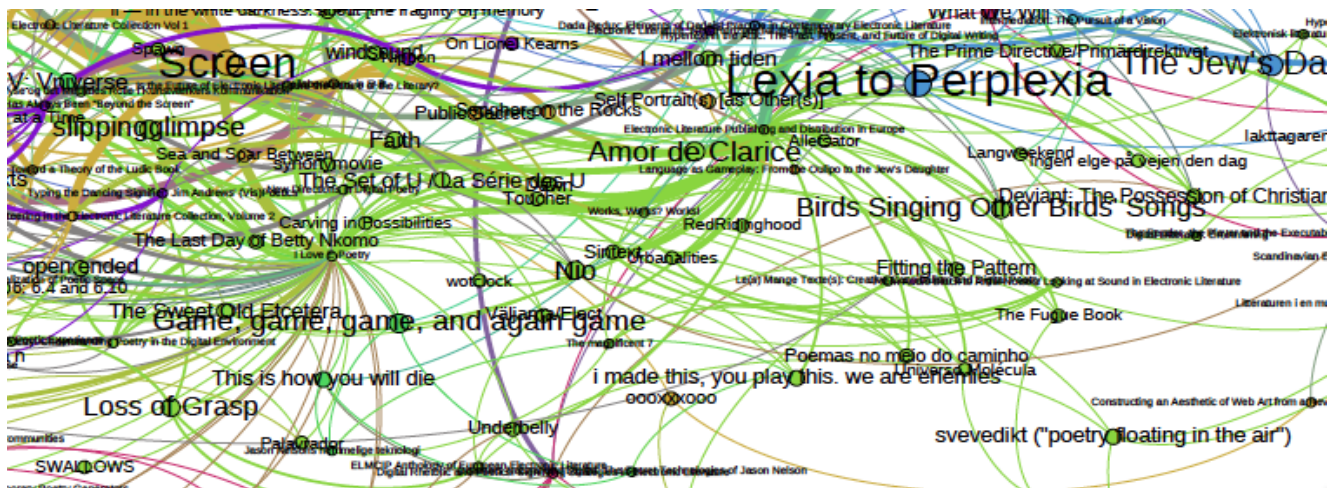
Fig. 16 Detail of visualization -- Flash, kinetic poetry, innovative UI cluster.

Another type of cluster evident are those that form around a small set of related works. The "Taroko Gorge" cluster is an example of this. The set includes some of the many different poetry generators based on remixes and over-writings of Nick Montfort`s original javascript source code. A recent burst of critical writing on the series of poems has resulted in the formation of a tightly interlinked cluster in the visualization.
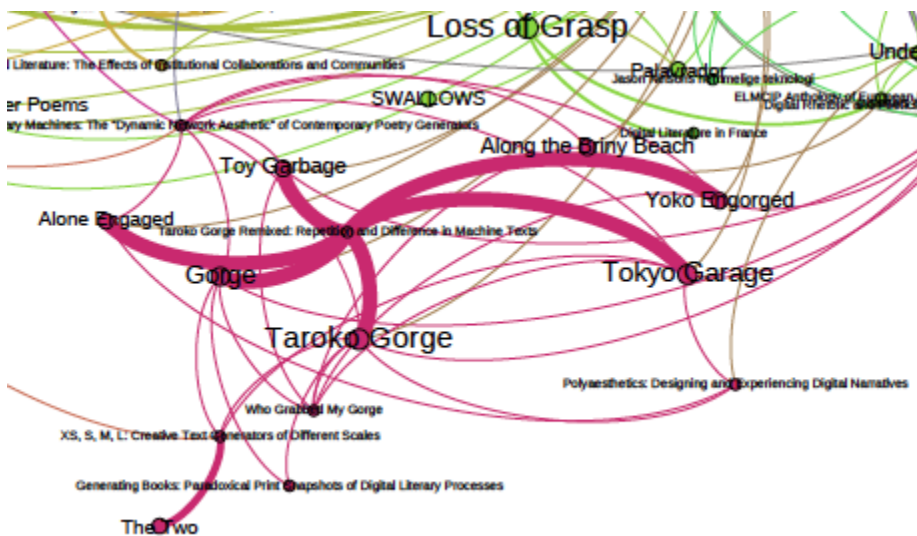

Fig. 17 Detail of visualization -- The "Taroko Gorge" nebula.

Though I have not gone into great detail in the analysis of network community structure, I think we can begin to see the potential usefulness of modularity analysis of the reference dataset. The technique might both verify and provide support for categories we already intuit, and show us groupings and patterns of generic or thematic association we would not have spotted before seeing them in the visualization.

The set of creative works with four or more citations in the Knowledge Base still includes 170 items. To reduce the set further, let's reduce the list to those works with eight or more citations in the set and come back to our earlier consideration of an "emerging canon." Although I don't think we could comfortably call this the canon, as a thought experiment it might be useful to ask "what if it were the canon?" and take a slightly closer look at the set to see what it would say about the field of electronic literature as a whole.

Table of Most-Cited Works: Works Referenced by 8+ Items of Critical Writing in the ELMCIP Electronic Literature Knowledge Base as of July 6, 2013

Title	Year	Author	# References
afternoon, a story	1990	Michael Joyce	76
Patchwork Girl	1995	Shelley Jackson	57
Victory Garden	1991	Stuart Moulthrop	37
The Unknown	1999	William Gillespie, Scott Rettberg, Dirk Stratton, Frank Marquardt	21
Façade	2005	Michael Mateas, Andrew Stern	21
Lexia to Perplexia	2000	Talan Memmott	19
Screen	2003	Noah Wardrip-Fruin, Josh Carroll, Robert Coover, Shawn Greenlee, Andrew McClain, Benjamin Shine	17
Text Rain	1999	Romy Achituv, Camille Utterback	16
Califia	2000	M. D. Coverley	16
ELIZA	1966	Joseph Weizenbaum	15
The Dreamlife of Letters	2000	Brian Kim Stefans	14
The Impermanence Agent	1999	Noah Wardrip-Fruin, Adam Chapman, Brion Moss	13
Colossal Cave Adventure	1976	Will Crowther, Don Woods	13
Cent mille milliards de poèmes	1961	Raymond Queneau	13

The Jew´s Daughter	2000	Judd Morrissey	13
Twelve Blue	1996	Michael Joyce	12
Uncle Buddy's Phantom Funhouse	1992	John McDaid	12
Game, game, game, and again game	2007	Jason Nelson	11
Amor de Clarice	2005	Rui Torres	11
Arteroids	2003	Jim Andrews	11
Hegirascope	1995	Stuart Moulthrop	11
Zeit für die Bombe	1997	Susanne Berkenheger	10
88 Constellations for Wittgenstein (to be Played with the Left Hand)	2009	David Clark	10
34 North 118 West	2002	Jeff Knowlton, Naomi Spellman, Jeremy Hight	10
Nio	2001	Jim Andrews	10
Birds Singing Other Birds' Songs	2001	Marîa Mencîa	10
Blue Company	2002	Rob Wittig	10
These Waves of Girls: A Hypermedia Novella	2001	Caitlin Fisher	10
Loss of Grasp	2010	Serge Bouchardon, Vincent Volckaert	10
slippingglimpse	2006	Stephanie Strickland, Cynthia Lawson Jaramillo, Paul Ryan	10
worm applepie for doehl	1997	Johannes Auer	9
Dakota	2001	Young-Hae Chang Heavy Industries, Young-Hae Chang, Marc Voge	9
my body & a Wunderkammer	1997	Shelley Jackson	9
Taroko Gorge	2009	Nick Montfort	9
in absentia	2008	J. R. Carpenter	9
Translation	2004	John Cayley	9
TOC: A New-Media Novel	2009	Steve Tomasula	9
riverIsland	2007	John Cayley	8

Tokyo Garage	2009	Scott Rettberg	8
V: Vniverse	2002	Stephanie Strickland, Cynthia Lawson Jaramillo	8
White-Faced Bromeliads on 20 Hectares	1999	Loss Pequeño Glazier	8
ppg256 (Perl Poetry Generator in 256 Characters)	2008	Nick Montfort	8
Zork 1: The Great Underground Empire	1981	Marc Blank, David Lebling	8
Faith	2002	Robert Kendall	8
Twilight, A Symphony	1996	Michael Joyce	8

This list will change as the coverage of the Knowledge Base becomes more complete. In spite of the language imbalance of the dataset made clear by the dominance of the English-language works in this list, from some other perspectives it is a remarkably balanced set of work, and the list also includes some informative surprises. The set includes a good temporal distribution of works, including classics and antecedents (in the case of Queneau's *Cent mille milliards de poèmes*) as well as more contemporary work. The greatest concentrations of works in the list are those published between 1999-2002.

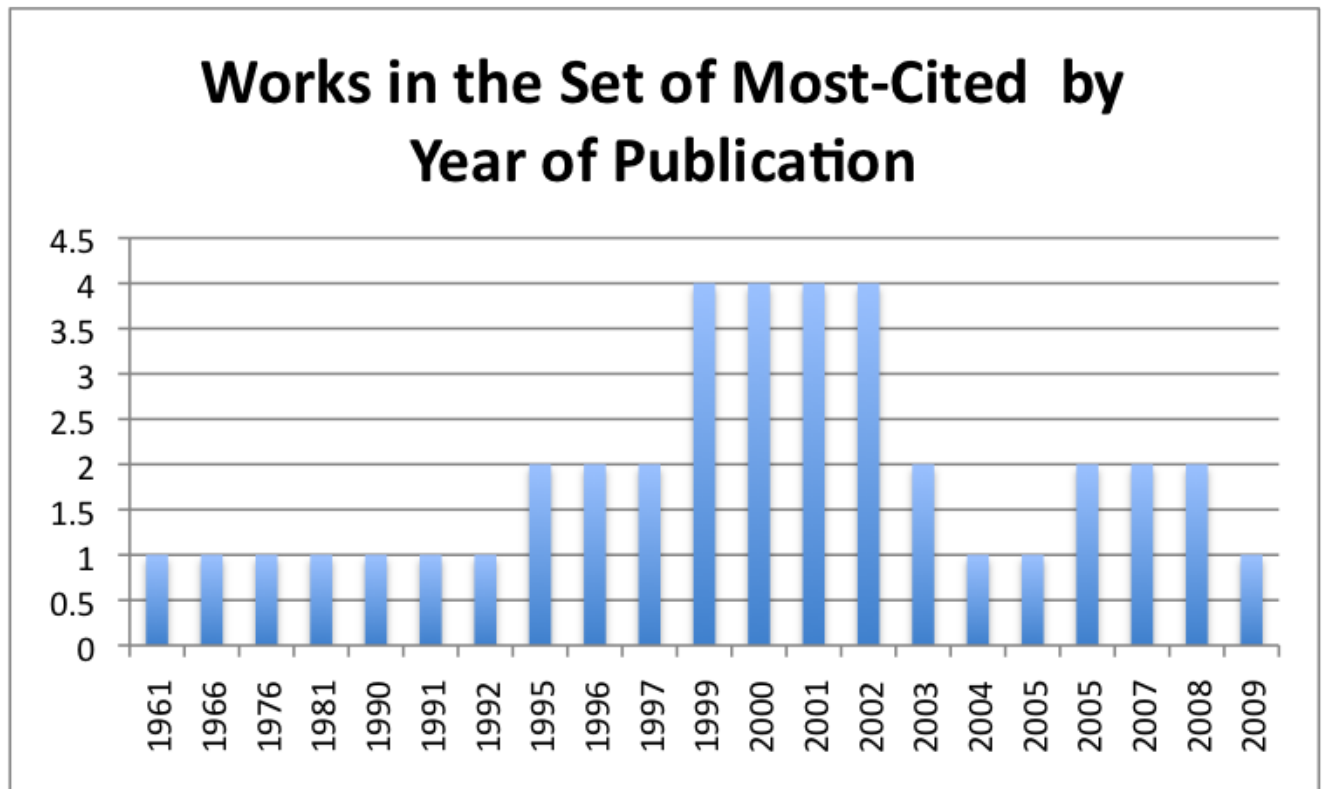## Works in the Set of Most-Cited by Year of Publication

Fig. 18 Distribution of creative works in the set cited by 8+ articles of critical writing by year of work's publication.

Though English language and particularly American works dominate the set, German, French, and Portuguese language works are also present on the list.

Among the genres of work represented on the list are hypertext fiction (both early works published on CD and published on the Web), literary text installations and CAVE works, ludic works that involve the conventions of games, kinetic poetry, interactive fiction, interactive drama, email narrative, visual poetry and works that reference the concrete poetry tradition, works that harvest and integrate texts from the web, poetry generators, a locative narrative, and works that emphasize aspects of user interaction. The list is slightly weighted towards narrative, though there are some strong representatives of poetic practice. One of the surprises in the set may be that it includes a number of works that are frequently cited as boundary or borderline examples of electronic literature in which interactivity and visual effect are dominant aspects, such as Jim Andrews's *Nio*, Maria Mencia`s *Birds Singing Other Birds' Songs*, and Romy Achituv and Camille Utterback's *Text Rain*.

From a pedagogical perspective, this list could serve as an excellent starting point for a syllabus, for instance in an introductory course on electronic literature, as it includes a wide variety of genres and technical modalities of electronic literature and a range of exemplars including early and recent work. Because the works are included here on the basis of their frequent citation, there would also be an ample set of secondary sources to include in the course, many of them available online and linked directly from the Knowledge Base. Although there would be accessibility issues with a few of the works in the set, notably some of those published by Eastgate, the majority of the works are still accessible and those that are not are well-enough documented in the critical literature of the field that they could still be integrated in a fashion into the curriculum.

The following visualization represents the tags on the works in the set.
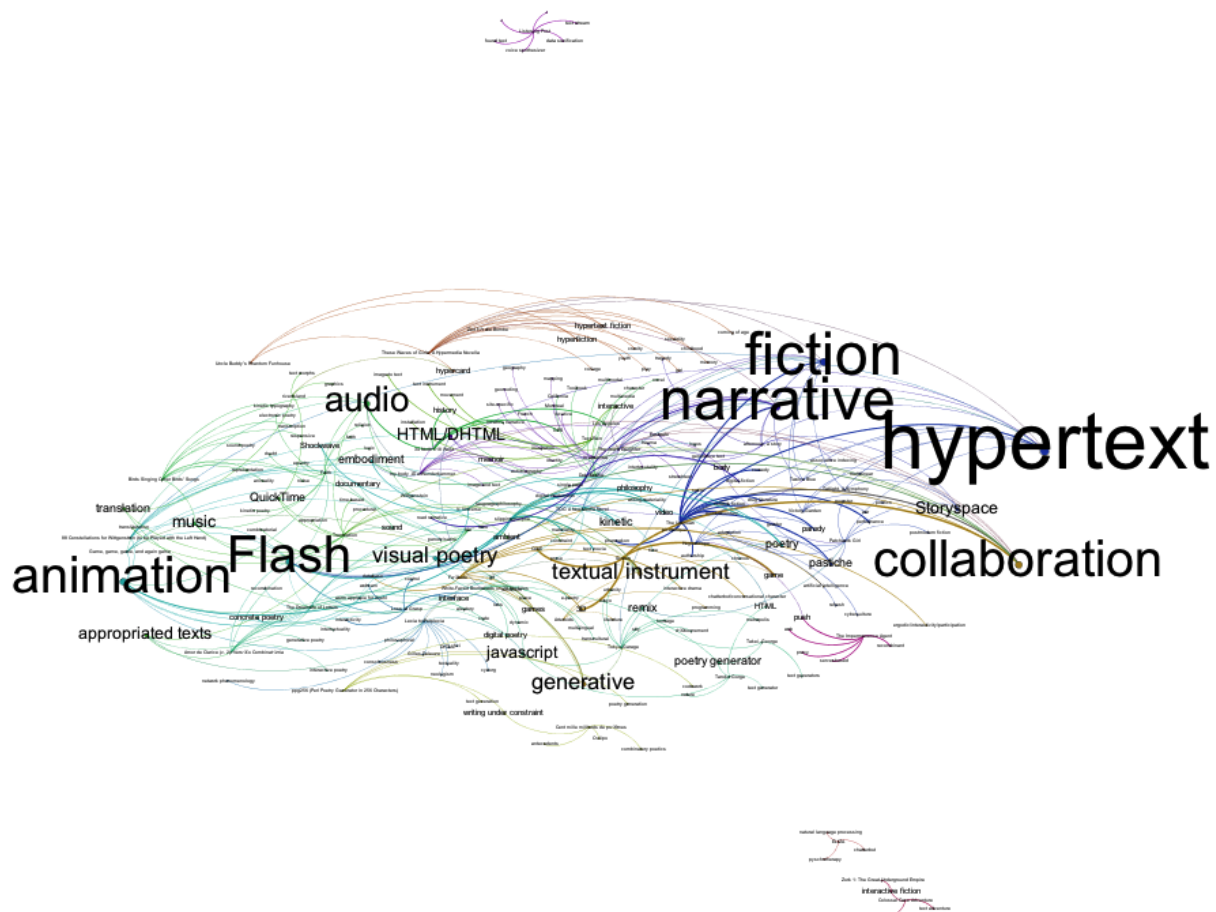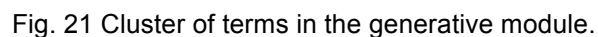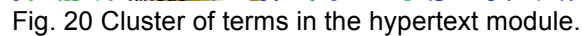

Fig. 19. Visualization of the tags used on a set of 46 works with 8 or more citations in critical writing.

Any analysis of works based on their tags in the ELMCIP Electronic Literature Knowledge Base should be understood as partial, biased, and contingent. Tags are an unstructured, contributory vocabulary. The individual contributor adding or editing the record chooses the tags. The system attempts to autocomplete using tags that are already in the system as the contributor types, but if a given tag has not been used before, it will be added to the vocabulary. So in comparison to the more semantically structured relations in the database, tags are a folksonomy and something of a "wild west" one at that. At our development meetings for the Knowledge Base, what to do with the tag system is still an open question. We periodically try to do some "weeding" to remove tags that are used in particularly idiosyncratic ways or to combine obvious synonyms. As long as the vocabulary is open and unstructured, it will however be unevenly applied, resulting for instance in different tags that mean the same thing, some works that are very minimally tagged while others are tagged extensively, etc. On the other hand, we are reluctant to over-determine the tagging system by establishing a rigidly structured set of fixed categories: our logic is that tags is where new concepts, methods, and genres can "bubble up" in the collective consciousness represented by the database. A very complex structured vocabulary would also require users who are trained to utilize it. This works well for some other projects: the records in NT2's Repertoire des arts ets littératures hypermédiatiques for example have been developed only by researchers and research assistants trained in the structured ontology they use to tag aspects of the work, such as interactivity, media format, content (theme/genre), and platforms utilized. In principle, we want to maintain a lower learning curve for contributors to the ELMCIP Knowledge Base: any author or scholar working in the field should be able to act as a contributor without a great deal of training. The approach to tagging in the ELMCIP Knowledge Base is still evolving: we want to preserve the ease of use and openness of the current system, but move some of the functions of tagging to more structured mechanisms. In 2013, for example, we will implement a Software/Platform content type, for records that describe, include version histories, and provide attribution to developers of the tools and environments used to make electronic literature. The works will then be cross-referenced directly to the record for the software, and tags will no longer be used for this purpose.

Because the tags are entered individually and different contributors have different concerns and methods in tagging, the list of tags used in the set is predictably diverse. Only ten tags are used more than five times in the set, and the vast majority of the terms are used only once. We can tell that the set is unevenly tagged, for example by the fact that the term "poetry" is used only three times, while there are many more poetry-based works in the set. The top 10 terms—

including hypertext, fiction, narrative, collaboration, animation, Flash, audio, generative, and visual poetry—are primarily indicative of genre, platform, or media, while the terms used only once in the set provide a rich profile of more specific themes and methods. Together they provide some insight into the vocabulary of the field.

Tags used in work records for the 46 most-cited works: hypertext (12), fiction (10), narrative (10), collaboration (9), animation (9), Flash (9), audio (7), generative (5), textual instrument (5), visual poetry (5), Storyspace (4), appropriated texts (4), javascript (4), HTML/DHTML (4), music (4), remix (3), pastiche (3), embodiment (3), kinetic (3), poetry generator (3), QuickTime (3), poetry (3), translation (3), body (2), parody (2), 3D (2), games (2), interface (2), game (2), interactive (2), history (2), concrete poetry (2), ambient (2), push (2), interactive fiction (2), writing under constraint (2), hyperfiction (2), hypertext fiction (2), hypercard (2), digital poetry (2), sound (2), HTML (2), documentary (2), philosophy (2), Shockwave (2), video (2), memoir (2), tragedy (1), associative indexing (1), Eastgate (1), trauma (1), adaptation (1), postmodern fiction (1), gender (1), monster (1), war (1), politics (1), multilinear (1), authorship (1), performance (1), satire (1), comedy (1), picaresque (1), network fiction (1), pharmakon (1), drug literature (1), mockumentary (1), database (1), interactive drama (1), artificial intelligence (1), chatterbot/conversational character (1), ergodic/interactivity/participation (1), textuality (1), cyborg (1), hci (1), consciousness (1), network phenomenology (1), philosophical (1), DHTML (1), Gilles Deleuze (1), neologism (1), CAVE (1), installation (1), text instrument (1), movement (1), image to text (1), novel (1), multimedia (1), California (1), character (1), multimodal (1), Toolbook (1), chatterbot (1), psychotherapy (1), natural language processing (1), feminism (1), recombination (1), recombinant (1), web (1), server-based (1), proxy (1), text adventure (1), 1970s (1), combinatory poetics (1), Oulipo (1), antecedents (1), generative text (1), intertexutality (1), single page (1), stretchtext (1), retro (1), parody/satire (1), aleatory (1), code (1), generative poetry (1), interactive poetry (1), intertextuality (1), literature (1), programming (1), refresh (1), cyberculture (1), found text (1), networked audiovisual installation (1), voice synthesizer (1), data sonification (1), web harvested work (1), text stream (1), language philosophy (1), Wittgenstein (1), locative narrative (1), GPS (1), Los Angeles (1), combinatorial (1), noise (1), sound poetry (1), animality (1), transcription (1), kinetic typography (1), representation (1), responsive (1), girl (1), coming of age (1), sexuality (1), play (1), cruelty (1), collage (1), childhood (1), youth (1), memory (1), interactivity (1), webcam (1), control (1), loss (1), materiality (1), writing-materiality (1), text movie (1), gif (1), digital modernism (1), digital fiction (1), road narrative (1), autobiography (1), image and text (1), identity (1), illustration (1), text generator (1), nature (1), text generators (1), French (1), mapping (1), locative (1), geography (1), geocoding (1), Montreal (1), site-specific (1), constraint (1), procedural (1), transliteration (1), appropriation (1), time-based (1), time (1), prose (1), chronos (1), logos (1), text morphs (1), graphics (1), city (1), urbanity (1), metropolis (1), tokyo (1), homage (1), détournement (1), place (1), e-poetry (1), dynamic (1), multilingual (1), transcultural (1), poetry generation (1), text generation (1), codework (1), epoetry (1), electronic poetry (1), kinetic poetry (1), religion (1), faith (1), doubt (1), logic (1).
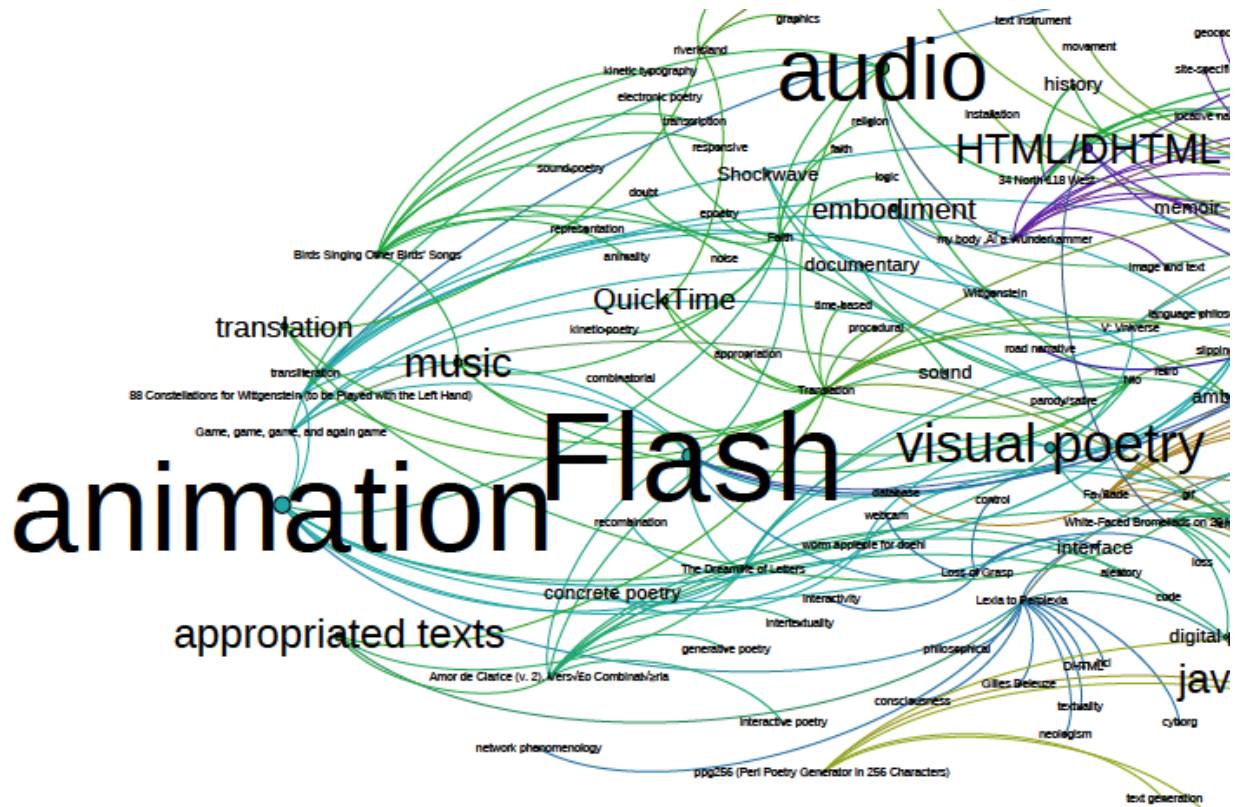
Weaknesses of method aside, the extant tags nevertheless tell a story about the field of electronic literature and its concerns. The network modularity of the visualization also reveals connections of genres and platforms, in particular the two anchoring terms on the right and left sides the visualization, with animation, Flash, and visual poetry dominant terms on the left axis, and hypertext, narrative, fiction, and Storyspace on the right, establishing a kind of spectrum of practices that have been dominant in the discourse of the field.

Fig. 20 Cluster of terms in the hypertext module.

Fig. 21 Cluster of terms in the generative module.

Fig. 22 Cluster of terms in the animation / Flash module.

The clustering of terms opens up some interesting questions, for instance about the relationship between platform and literary genre in the field of electronic literature, which should be a rich basis for further research. For example we might ask if the hypertext authoring systems of the 1990s were particularly well-suited to narrative, while the period in which many e-lit authors and developers in the 2000s were drawn to Flash led to a dominance of forms and genres of e-lit related to visual poetry and animation?

Conclusion

The results I have presented here are clearly preliminary. At this point we can say that while the ELMCIP Electronic Literature Knowledge Base offers the most extensive documentation of citation of creative works of e-lit by critical writing available, the record is still woefully incomplete. More work is clearly necessary in developing this historical record of critical activity, and if it is to be more accurate, it will need to be done by a larger group of contributors to the Knowledge Base itself. The small group of researchers in the University of Bergen Electronic Literature Research Group cannot provide this documentation alone. An accurate mapping of the field of electronic literature will require the participation of the international community of electronic literature scholars. The contributory model of the Knowledge Base is such that much

broader participation in developing—in crowdsourcing—the record of critical references to creative works is possible. As more authors of creative works and authors of critical writing make documenting their work in the shared research environment of the database an aspect of their workflow, the representation of the field and what we might or might not want to refer to as "the canon" of electronic literature will become more accurate.

To circle back to the question of whether or not we would benefit from having an established "canon" of electronic literature, my opinion is that this will be a matter of enduring and unresolved debate. The most likely outcome will be that we abandon the language of "canonicity" with all its rhetorical and historical baggage and instead focus on what is useful about having a greater understanding of a shared set of references within a field of scholarly and creative practices. Electronic literature is a field for various reasons prone to a sense of "presentism." Because of its intrinsic relationship to the technological apparatus in which it is produced, read, and processed, novelty is itself often cited as a literary value. While I don't deny that innovation and experimentalism often provide compelling reasons to engage with a creative work, if we place novelty at the apex of a hierarchy of values, we risk blinding ourselves to the work that preceded any given innovation, and thus operating without any sense of history. We also risk focusing so intently on technological innovation that we might lose sight of what any given work is actually *about*; and content matters in literary work.

This type of analysis can be one element in establishing a deeper sense of historical context: enabling us to not only see "what's happening now?" but "what has happened and how does what's happening now relate to what happened then?" One consequence of the documentation project of the Knowledge Base has for example been the recovery of some of some lost works. As we dig more deeply into the criticism of e-lit that took place in the 1990s, we are able to locate and recover some early works of web-based e-lit from the Internet Archive. So engaging with the critical record in this way can actually have the effect of revivifying works that were written about at one time but not written about enough to continue circulating within the critical discourse: that is to say actually bringing them back from the void into which they were by circumstance cast.

Some of the concerns with the rhetoric of canonization in print literature are less pressing in considerations of e-lit. There is little evidence, for example, that scholars working in the field have felt restricted to teaching any common set of works: indeed the material conditions of

preserving digital artifacts are such that many literary projects are likely to be inoperable before they are widely circulated within reading communities on a sustained basis. Even the two most-cited works of electronic literature, *afternoon, a story* and *Patchwork Girl* can at present be difficult to run on many contemporary operating systems. So there is little reason to fear that having a common set of references to earlier works and understanding their critical reception will crowd out or marginalize new works that can be experienced on current operating systems.

There are two audiences in particular that I think are well served by having access to a historical list of frequently cited works of electronic literature. The first audience is a pedagogical one: teachers who may for example be teaching a survey of electronic literature for the first time will find on the above list of most-cited works the initial basis for a curriculum of both primary and secondary sources. This might serve to lower some of the boundaries for entry to the discourse of the field. The second audience served by a historical understanding of criticism of electronic literature might be that of humanities, even digital humanities, scholars who consider the practice and scholarship of electronic literature to be *dubious*. I am certain that any of us who have been operating in this field for a sustained period can recount some instances of encountering the blank, glossy-eyed stares of Victorianist, or Modernist, or Comp Lit, or Creative Writing colleagues who insist on asking the same questions over and over again: "What a *novel* approach to literary studies. But is it *really real*? It appears not to be about *books*. Isn't it just a *passing fad*?" These misconceptions have very real effects, for example in curriculum, tenure, and promotion committees and in the situation and placement of the field within the academic context. In response, we can now offer skeptics solid evidence of thirty years of sustained critical activity (including a great deal of printed, as well as digital, matter) that has constituted a dynamic field. Whether or not you call it an emerging canon, it is certainly not an apparition.

I have not dealt in great detail here with debates in digital humanities surrounding macroanalytic "distant reading" methods and information visualization applied to literary materials. As this paper is a product of such methods, I hope that it serves as an implicit argument for their use. But let me be clear about my own position: I don't think that distant reading is any sense a replacement for closer readings which engage very directly with the language, technical apparatus, interface, visual presentation, interactive semiotics, etc. of any individual digital literary artifact. Nor do I think it is a replace for "middle reading" that examines works within the contexts and conditions of a particular genre, community of practice, or historical context. I

argue instead that macro-level readings of literary communities and content should be regarded as one level of zoom, which can range from reading an entire field distantly to reading a single line of code exhaustively.

The papers presented by members of the University of Bergen Electronic Literature Research Group at the 2013 Chercher le Texte Electronic Literature Organization conference should be regarded as first steps towards integrating macroanalytic research into the tool-kit of electronic literature criticism. Databases such as the ELMCIP Electronic Literature Knowledge Base, the ELO's Electronic Literature Directory, the NT2 Repertoire des Arts et Littératures Hypermédiatiques, the Po.Ex Archive of Experimental Portuguese Literature, the Brown University Digital Literary Arts Archive, and others are developing into extremely rich sources of metadata about creative and critical writing in the international field. I hope that our initial forays into this mode of research and these digital methods provide some evidence of potential ways to use these databases not only as bibliographic reference, documentation, and archival resources, but also as the basis for new modes of research in their own right. We hope to produce an edited collection of essays focused on using visualization research methods to analyze the field of electronic literature within the next year.

Acknowledgements

Base. Our mutual research in this area is a frequent topic of discussion both in the office and at the dinner table of our home in Bergen.

References

Ensslin, Astrid. 2007. *Canonizing Hypertext: Explorations and Constructions*. New York: Continuum.

Rettberg, Jill Walker. 2013. "An Analysis of Dissertations About Electronic Literature." ELO 2013 Conference: Chercher le Texte, Paris.

Rettberg, Jill Walker and Scott Rettberg. 2013. "Mining the Knowledge Base: Exploring Methodologies for Analysing the Field of Electronic Literature." Digital Methods Winter Workshop and Mini-Conference, University of Amsterdam.

Rettberg, Scott, with Eric Dean Rasmussen. Forthcoming 2013. "The ELMCIP Electronic Literature Knowledge Base: Project Report" in Electronic Literature as a Model of Creativity and Innovation in Practice. Bergen / Morgantown: ELMCIP and West Virginia University Press.

Saemmer, Alexandra. 2012. "Evaluating Digital Literature: Social Networks, Selection Processes, and Criteria" in *Remediating the Social.* Bergen: ELMCIP. 83-87.

Creative Works Referenced
Works specifically referenced in the body of the text.

Achituv, Romy and Camille Utterback. 1999. *Text Rain*. ELMCIP record:: http://elmcip.net/node/467

Andrews, Jim. 2001. *Nio*. ELMCIP record: http://elmcip.net/node/817

Blanc, Marc and David Leibling. *Zork 1: The Great Underground Empire*. ELMCIP record: http://elmcip.net/node/3876

Coverley, M.D. *Califia*. 2000. Watertown: Eastgate. ELMCIP record: http://elmcip.net/node/723

Crowther, Will and Don Woods. 1976. *Colossal Cave Adventure*. ELMCIP record: http://elmcip.net/node/593

Gillespie, William, Scott Rettberg, Frank Marquardt, and Dirk Stratton. 1998. *The Unknown*. ELMCIP record: http://elmcip.net/node/662

Jackson, Shelley. 1995. *Patchwork Girl*. Watertown: Eastgate. ELMCIP record: http://elmcip.net/node/239

Joyce, Michael. 1990. *afternoon, a story*. Watertown: Eastgate. ELMCIP record: http://elmcip.net/node/236

Mateas, Michael, and Andrew Stern. 2005. *Façade*. ELMCIP record: http://elmcip.net/node/1293

Memmott, Talan. 2000. *Lexia to Perplexia*. ELMCIP record: http://elmcip.net/node/65

Mencia, Maria. *Birds Singing Other Birds' Songs*. ELMCIP record: http://elmcip.net/node/582

Montfort, Nick. 2009. Taroko Gorge. ELMCIP record: http://elmcip.net/node/3595

Morrissey, Judd. 2000. *The Jew's Daughter*. ELMCIP record: http://elmcip.net/node/78

Queneau, Raymond. 1961. *Cent mille milliards de poèmes*. ELMCIP record: http://elmcip.net/node/4694

Weizenbaum, Joseph. 1966. ELIZA. ELMCIP record: http://elmcip.net/node/1539