

# A Network Analysis of Dissertations About Electronic Literature

Jill Walker Rettberg

Paper for ELO Paris 2013. Part of the [ELMCIP panel](#). Current draft as of August 20, 2013.

## **Abstract**

*More than 60 dissertations in the field of electronic literature have been documented in the ELMCIP Electronic Literature Knowledge Base, including tags, abstracts and in most cases links to full texts of the dissertations. This paper performs a network analysis of the citations in 29 of these dissertations to identify trends, patterns and information about an emerging canon.*

## **Introduction**

The first dissertation about electronic literature was written in 1976 by James Richard Meehan, and was titled *The Metanovel: Writing Stories by Computer*. Since then, at least 60 more dissertations have been written on electronic literature. This vast corpus of research literature about electronic literature has not previously been studied as a whole.

This paper begins to map creative works in electronic literature as represented by dissertation on electronic literature. Information about references to creative works in the dissertations has been entered into the ELMCIP Electronic Literature Knowledge Base, and keywords and tags for each dissertation were also registered. The data was then imported into the open source network analysis software Gephi, where it was visualised and analysed.

## **Changing topics, shifting concerns**

Nobody familiar with the growth of the field of electronic literature over the last years will be surprised to hear that more and more doctoral students are choosing to write their dissertations about electronic literature. In the first years, dissertations on electronic literature were occasional events, but in the last decade, there have consistently been several published each year, which allows for a critical mass of new scholars who can respond to each others work.

In this analysis my goal was not to select a representative sample but to analyse *all* dissertations on electronic literature, but without a doubt there are many dissertations missing from my data. Ideally we would have data about and the full text of every PhD dissertation written at any university in any country about electronic literature. In practice, the Electronic Literature Research Group at the University of Bergen has done its best to find as many dissertations as possible. We have searched archives such as ProQuest and Google Scholar for “electronic literature”, but just as importantly we have made a point of entering information about every dissertation we came across, saw referenced or read a mention of. Our goal has not been Fortunately all the data is documented in the ELMCIP Electronic Literature Knowledge Base, and as new dissertations are recorded in the knowledge base it will be simple for us or other researchers to re-run our analysis with the larger corpus.

Most dissertations are documented online in institutional repositories. Usually at least an

abstract is available, often the author has also provided tags or keywords, and in many cases the full text of the dissertation is available. We also added tags when these were not given by the author, and based this on the abstract, and the full text if we had access to it. When we could not find the full text online we attempted to contact the author to ask for a copy. There are, however, some dissertations we could not access beyond the title and abstract. The Knowledge Base allows cross-referencing between nodes, so we have added cross-references (links) from the entry about each dissertation to the creative works it discusses, although this was not possible for dissertations we did not have full text access to. A few dissertations have been left out because we could not find a person able to read the language of the dissertation or the works referenced well enough to register the data - unfortunately, Xiaomeng Lang's work on Chinese electronic literature (Lang 2008) is one of these, and Asian electronic literature is sorely underrepresented in the Knowledge Base. We hope to be able to include such material in future analyses.

Simple word clouds of the tags used to describe the dissertations clearly show a shifting focus in the study of electronic literature over the years. The first dissertation on electronic literature was James Meehan's "The Metanovel: Writing Stories by Computer" (1976), followed by Mary Ann Buckles' "Interactive Fiction: The Computer Storygame 'Adventure'" (1985) and then a steady stream from 1991 onwards.

I split the dissertations into three groups: 1976-2001, 2003-2007 and 2008-2012. We found no dissertations on electronic literature published in 2002. The first time period shows a clear focus on technology and on the mechanics of reading these kinds of text, as shown in the tag clouds in Figure 2. The tag clouds are generated from tags assigned to the dissertations in the Knowledge Base. When the dissertation author specified keywords on the dissertation we used these, and in other cases we added them based on the description and content of the dissertation. The tag clouds were generated using an online tool that displays more frequently used tags in a bigger font. Colours do not signify anything in these images.

As you can see in Figure 1, Hypertext is by far the most common tag (10 of 15 dissertations) and the next most popular tags are only used in three dissertations each (fiction, text, reading, computer) followed by five tags used in two dissertations each (computational, interactive, new media, theory, cybertext).



Figure 1: The tag clouds show the shifting focus in dissertations on electronic literature. From left to right, we see tags used to describe dissertations published between 1976-2001, in the centre dissertations published between 2003-2007 and to the right dissertations published between 2008-2012.

The next batch of dissertations was published from 2003 until the end of 2007. We see hypertext is still a very important tag, but the focus on technology and the mechanical is dwarfed by more *literary* words: fiction, literature, narrative, as well as by words showing the broadening of the field: game, art, media.

The third group of dissertations covers the period from 2008-2012, and the most obvious shift is the way “hypertext” has disappeared and “digital” has taken its part. “Poetry” has grown a lot, and “fiction” is barely mentioned. the more general term “literature” has shrunk a little, and we also see how “media” has grown steadily over the three periods. The last few years have seen several dissertations on quite specific genres: interactive fiction, kinetic poetry and generative texts, and this is reflected in the more genre-focused vocabulary of current dissertations.

### **Diversity is the rule**

After going through the dissertations [28 in this draft, close to 60 when completed] and adding links in the Knowledge Base to the creative works they referenced, I was able to see clear patterns in the ways that the dissertations discuss creative works. The most important finding is that PhD dissertations about electronic literature reference a great diversity of creative works.

As of early July 2013, references to creative works are added for 29 of the 33 dissertations published from 2006-2013 that we have records for in the Knowledge Base. The remaining five dissertations were not included either because we lacked access to the full text, or, in the case of two dissertations, due to language challenges (Hungarian and Chinese). These could be added in later analysis.

A total of 341 separate works were referenced in the dissertations, and there was a lot less redundancy than might have expected. 265 creative works were only cited by one of the 28 dissertations. 57 were cited by two dissertations, and just 19 by three or more dissertations. On average, then, each dissertation cites 16 different creative works. In practice, of course, some cite just a few works while others cite many, but certainly most dissertations cite many creative works. This suggests that the claim that scholarship in electronic literature ignores the creative works and only focuses on theory is at least not the case in current dissertations.

This preliminary corpus gives us the following (preliminary) “top twenty” list of the most cited creative works in dissertations:

<b>Title of work</b>	<b>Citations</b>
afternoon, a story	9
Victory Garden	7
Patchwork Girl	7
Cent mille milliards de poèmes	6

Zork 1: The Great Underground Empire	6
The Impermanence Agent	5
ELIZA	5
Façade	5
my body--a Wunderkammer	4
Lexia to Perplexia	4
Deadline	4
The Last Day of Betty Nkomo	3
Sooth	3
Composition No. 1	3
Text Rain	3
The Legible City	3
The Unknown	3
The Policeman's Beard is Half-Constructed	3
Colossal Cave Adventure	3
JABBER: The Jabberwocky Engine	2

There are no great surprises in the list, these are all family works, although it's interesting to note that most of the works are early and well-known examples in their genres.

## Clusters of Works

Network analysis allows us to see connections between the dissertations and the works they discuss as a network graph. This kind of visualisation and analysis has its roots in social network analysis, a sociological methodology that started in the 1950s, and in network theory, a mathematical methodology allowing us to understand networks in, for instance, molecular structures, the spread of disease, airline traffic networks or mobile phone usage.

The connections between dissertations and the creative works they cite form a *bipartite* or *two mode network*, which means there are two types of nodes: dissertations and creative works. A reference from a dissertation to a creative work forms a connection or an *edge* in this network.

Figure 2 shows the complete network of dissertations and creative works that are referenced. Dissertations are shown in blue and creative works in red. Creative works have been sized according to their *indegree*, that is to say the number of inbound links or the number of different dissertations that reference them. The most-cited creative works tend to cluster in the middle of the network, and you can see other clusters around the edges. Just one dissertation in the sample, Anders Løvlie's 2011 *Textopia: Experiments with Locative Literature*, has no shared references with the others, and so it has floated out to the left with no connections to the main network. The blue spots with no connections are the dissertations that are in the Knowledge Base but that have not yet had references to creative works added.

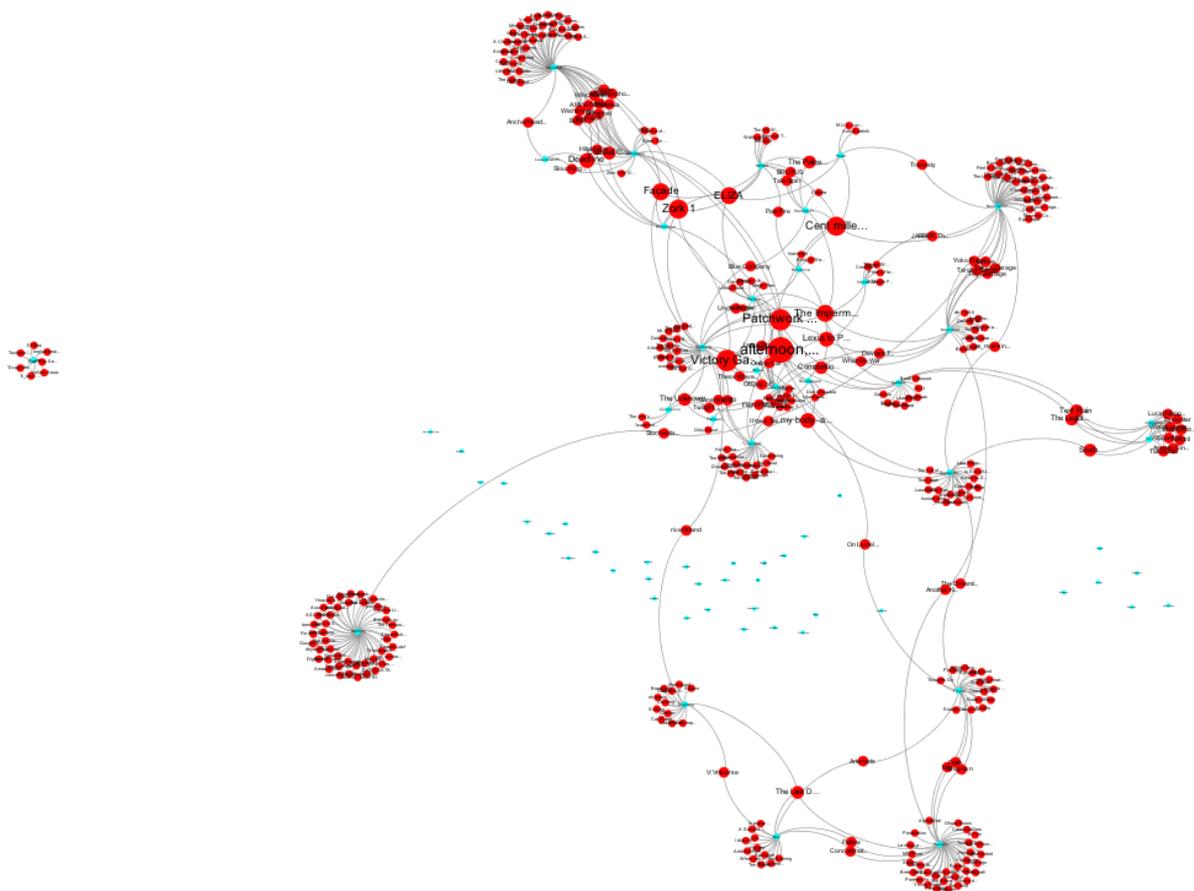


Figure 2: A bipartite or two-mode network of dissertations (in blue) and creative works (in red), sized by indegree. The layout was determined by the ForceAtlas 2 algorithm in Gephi, with Scaling (repulsion which determines the sparsity of the graph) set to 20 and Gravity (which pulls nodes towards the centre) set to 20. (*Dissertations-and-creative-works-28-diss-completed.php*)

Looking more closely at the graph, we can see clusters emerging around certain genres of electronic literature. For example, at the top of Figure 2 we see a cluster of dissertations about interactive fiction, shown in the detail view in Figure 3. Nick Montfort's 2007 dissertation *Generating Narrative Variation in Interactive Fiction* cites a fan of works that are not cited by other dissertations, but also many that are cited in Jeremy Douglass's 2007 dissertation *Command Lines: Aesthetics and Technique in Interactive Fiction and New Media*. Montfort has one shared reference, *Anchorhead*, with Van Leavenworth's 2010 dissertation *The Gothic in Contemporary Interactive Fictions*, which also has a number of shared references with Douglass' dissertation.

We see that *Façade* and *ELIZA*, which both centre on conversational characters, and *Zork*, the first commercial interactive fiction, are referenced by the three dissertations on interactive fiction, but that they are pulled away from it in the layout of the graph. That is because they are also referenced by many other dissertations that do not reference other interactive fictions. This is where the network visualisation really begins to get interesting, because it allows us to see how the different creative works relate to each other. *Façade* and *ELIZA* are in a *bridge* or *broker* position between dissertations on interactive fiction and dissertations on generative narrative and poetry, by Noah Wardrip-Fruin (2006), Fox Harrell (2007) and Daniel C. Howe (2009). This generative cluster is circled in green in Figure 3.

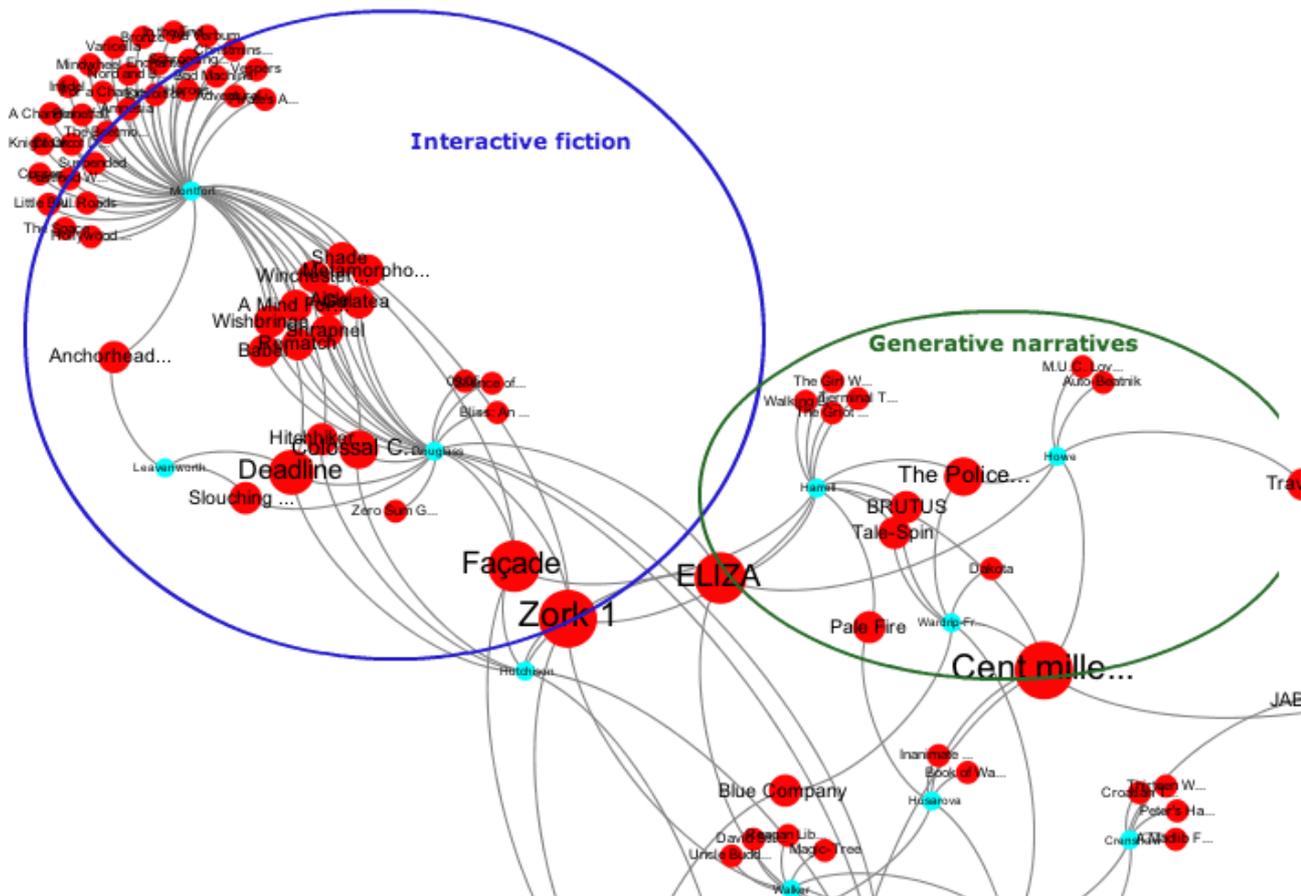


Figure 3: A detail of Figure 2 showing a cluster of dissertations on interactive fiction, circled in blue, and another on generative narratives, circled in green.



as suggested by Figure 2.

The portion of the network shown in Figure 4 might be most interesting as a way of navigating the field. Seeing the works laid out in this manner would be useful for newcomers to the field or anyone looking for new works to read and explore, as is particularly evident in the web-based browsable versions of networks like these presented in Scott Rettberg's paper at this conference (Scott Rettberg 2013).

## Relationships Between Creative Works

Another way of viewing the network is to convert it to a one-mode network: that is, to remove the dissertations from the network and instead view two creative works as being directly connected to each other if they are discussed in the same dissertation. Rather than saying that Engberg's, Naji's and Rosario's dissertations are connected to each other through their shared referencing of *The Last Day of Betty Nkomo*, a one-mode graph looking only at the creative works would instead say that *The Last Day of Betty Nkomo* is connected to

The process is engagingly described in Kieran Healy's

Using Jaroslav Kuchar's [Multimode Networks Transformations](#) plugin for Gephi I then converted the first bipartite network, consisting of dissertations connected to the creative works they referenced, into a monopartite or one-mode network consisting only of creative works. The original network had directed edges, because there was a clear direction of references: *from* a dissertation *to* a creative work. Kuchar's plugin requires undirected edges, so for the plugin to work, I exported the network as two csv files, one for the nodes and one for the edges, I changed the edge type to undirected in Excel, then imported both csv files into Gephi again and ran the plugin, asking it to remove the dissertation nodes.

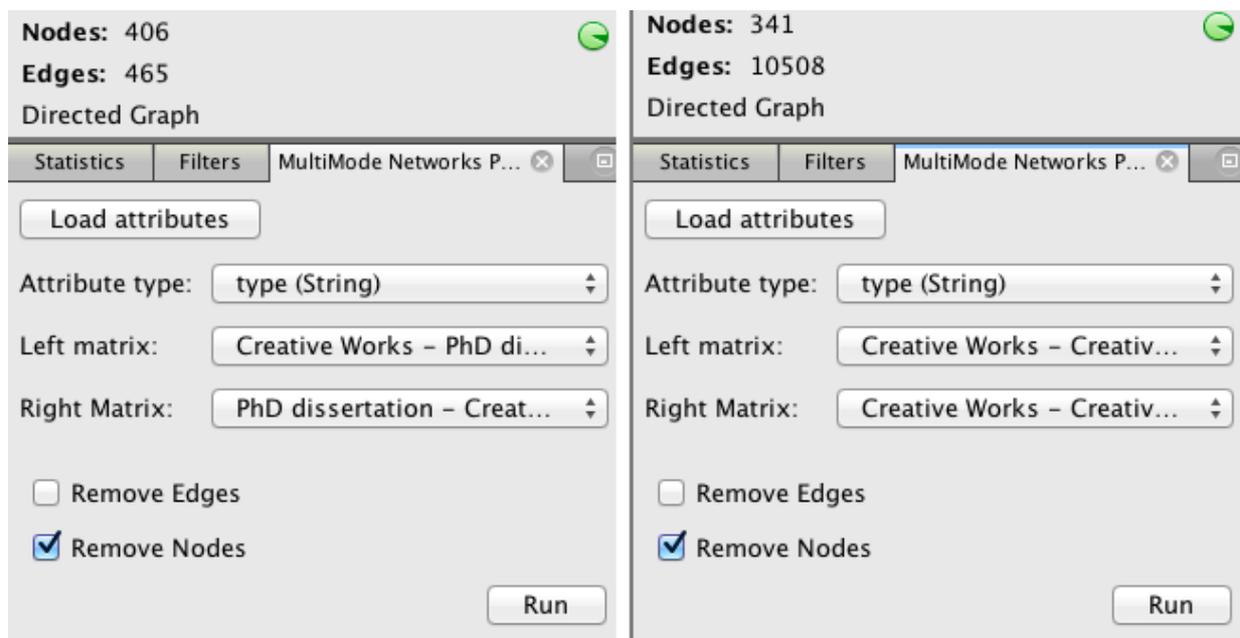


Figure 5: The Multimode Networks Transformation plugin interface. On the left, you see the interface as it was before the transformation. There are 406 nodes, including dissertations and creative works, and



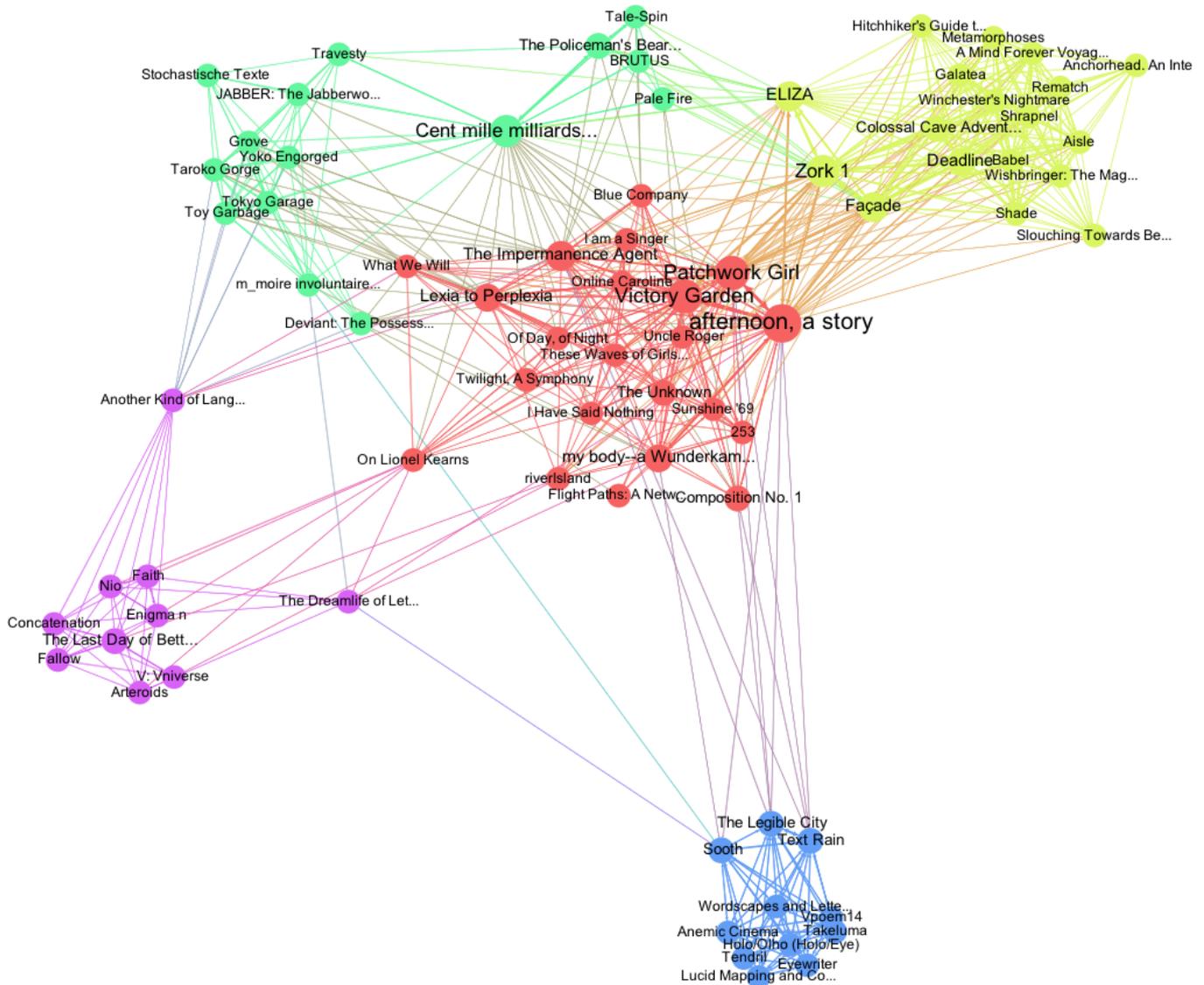


Figure 7: Creative works cited by at least two different dissertations, visualised as a one-mode network where an edge between two works means that they were cited in the same dissertation. Nodes are coloured by Gephi's modularity algorithm, and the layout is generated using the ForceAtlas 2 algorithm.

The genres in this graph are quite clear. The modularity algorithm groups works into just five modularity categories. At the top, in green, we see generative works, with the Oulipoan print classic *Cent mille milliards des poemes* an important hub in this group. On the upper right, interactive fiction and classic conversational characters form one cluster, in yellow. In the centre, in red, we see classic hypertext fictions and what we perhaps might call “mainstream” or “classic” works of electronic literature. The lower two clusters are interesting. They both represent more visual and kinetic poetry, but it’s not clear why they are as separate as they appear in this representation. Admittedly only three or four dissertations

have determined each of these two clusters, but it is surprising that they appear really quite cohesive and yet separate. The purple cluster can perhaps be thought of as kinetic poetry or digital poetry, whereas the blue cluster leans more towards the art world.

This representation of the field does leave us with some questions. There are individual works where the placement seems surprising. Why is Donna Leishman's *Deviant: The Possession of Christian Shaw* (2004) in the generative narratives group? Why is *Composition No 1.*, which is a printed novel in loose leaf format, in the lower right part of the red "classics" cluster, near the art poetry, rather than being more closely connected to the generative works? Perhaps the answers lie in the idiosyncrasies of the dissertation authors, rather than in any generic qualities of the works. And yet even with just 28 dissertations for our corpus, the overall graph corresponds well to real-world genres.

### **Ways of visualising a field**

Visualising the field of electronic literature by looking at creative works cited by dissertations provides one view of the field that is difficult to see by other means. The clusters do correspond to genres, unsurprisingly, since dissertation authors especially in the last few years tend to focus on a particular kind of electronic literature rather than try to write about it all, as was perhaps a greater temptation in the earlier days of the field.

An important advantage of using the Knowledge Base to visualise the field is that this makes it possible to include and even centre the analysis around creative works, which are not part of library catalogs or others systems that co-citation analysis might typically pick up.

In a data sprint at the Digital Methods Winter School at the University of Amsterdam this January, I worked with fellow scholars on alternative way of visualising the field of electronic literature: by feeding a selection of "seed books" in the field into the Amazon advertising API and retrieving books that are also bought by people who read the seed books. A full account may be read in Berry, Borra, Helmond, Plantin and Rettberg (forthcoming). Igt was immediately apparent that while our digital humanities seed books generated a fairly cohesive network that gave a fairly clear idea of what the digital humanities might be, the electronic literature seed books generated a far more disparate network.

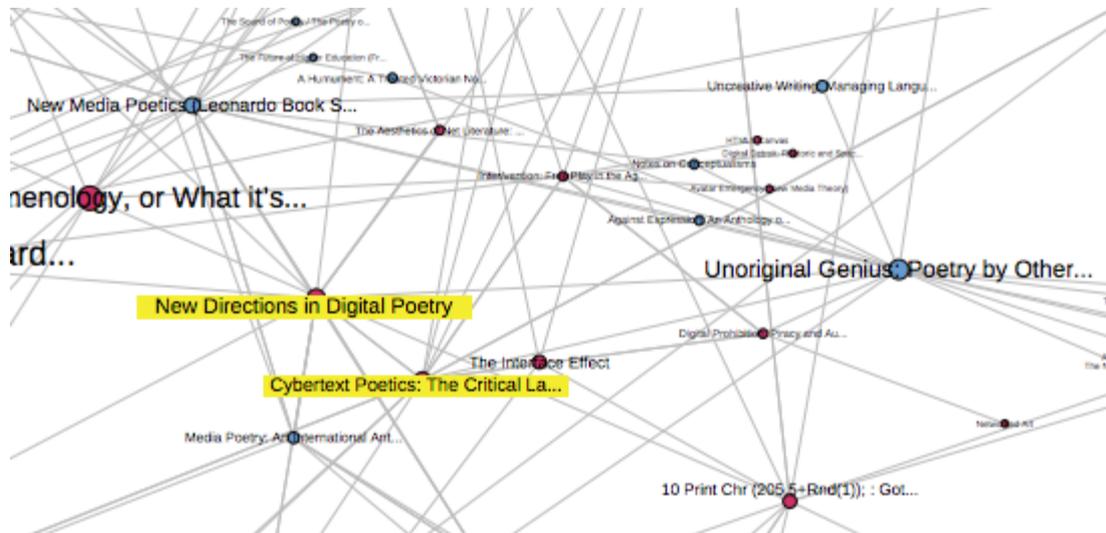


Figure 9: Detail of the electronic literature graph in Figure 8 showing a cluster of books on conceptual writing and digital poetics. The two titles marked in yellow were among the original seed books. Full image:

[https://wiki.digitalmethods.net/pub/Dmi/Winter13AmazonRecommendations/EL\\_FULL\\_FA2\\_ScaledIndex\\_ModularityColorCoding.pdf](https://wiki.digitalmethods.net/pub/Dmi/Winter13AmazonRecommendations/EL_FULL_FA2_ScaledIndex_ModularityColorCoding.pdf)

While the Amazon related books network for electronic literature does show interesting adjacent fields, such as the conceptual writing books in figure 9, a large and cohesive cluster on games studies, as well as a strong cluster on digital humanities, the lack of a strongly cohesive cluster of books specifically on electronic literature can be read as evidence that the field of electronic literature is not defined by books. There is little surprise in that, of course, though it is interesting that the digital humanities, which is also nominally centred around digital projects and methodologies, is far more clearly defined by its books.

The Amazon related books network does include three works of electronic literature: three of the early Storyspace classics are of course for sale on Amazon. People who buy Hayles' *Electronic Literature*, Landow's *Hypertext 3.0* or Murray's *Hamlet on the Holodeck*, apparently frequently also buy *afternoon, a story*, and people who buy *afternoon* often buy *Patchwork Girl* or *Victory Garden* - or techno-utopian or -dystopian books like Clay Shirky's *Here Comes Everybody* or Evgeny Morozov's *Net Delusion*.

As we detail in the paper discussing the Amazon related books networks, there are many possible methodological flaws and sources for error in the analysis, but it does serve as an interesting mirror view of the field of electronic literature that is, as the digital methods creed demands, completely based on data that is digitally native rather than curated by experts as the ELMCIP Knowledge Base is.

## Dissertations as Collective Curation

As shown by this and related research using network analysis to analyse data from the ELMCIP Electronic Literature Knowledge Base, there is much interesting work that can be done. The Knowledge Base offers a promising platform for scholars interested in

understanding the big picture of electronic literature, whether one is most interested in creative works or in scholarship about and around them.

In effect, this paper repurposes dissertations on electronic literature, using them as a form of collective curators of the creative works in the field. Individually, the scholars writing these dissertations were not intending their selection of a particular set of creative works to be appropriated by a scholar such as myself in order to map the field. Perhaps they would have chosen differently if they had realised that their work would be used in this way. If a scholar currently writing a dissertation on electronic literature hears or reads this paper, they may think differently about their selection. Perhaps the knowledge that their references will transform to links that can be used to visualise and understand a field will bring us full circles: if our references really do become links, then of course our collective scholarship should be viewed as a collective hypertext. Using visualisation software such as Gephi, we can make the map view of our collective field.

## Acknowledgements

I could not have done this work without the support and active participation of the University of Bergen Electronic Literature Research Group and of the ELMCIP project, both led by Scott Rettberg. While the analyses in this paper are my own, a lot of the data was entered by fellow research group members and by the many other contributors to the Knowledge Base. My switching use of “we” and “I” reflects this shared labour.

## References

Berry, David, Erik Borra, Anne Helmond, Jean-Christophe Plantin, Jill Walker Rettberg. Forthcoming. “The Data Sprint Approach: Exploring the field of Digital Humanities through Amazon’s Application Programming Interface.” Paper to be submitted for publication in September 2013. Data and early findings available at <https://sites.google.com/site/whatisdigitalhumanities/>

Hanneman, Robert A. and Mark Riddle. 2005. *Introduction to social network methods*. Riverside, CA: University of California, Riverside <http://faculty.ucr.edu/~hanneman/>

Healy, Kieran. 2013. “Using Metadata to find Paul Revere” *Kieranhealy.org* June 9. <http://kieranhealy.org/blog/archives/2013/06/09/using-metadata-to-find-paul-revere/>

Lang, Xiaomeng. 2008. *Der Dialog der Kultur und die Kultur des Dialogs: Die chinesische Netzliteratur*. PhD dissertation, University of Siegen. [http://dokumentix.ub.uni-siegen.de/opus/frontdoor.php?source\\_opus=398&la=de](http://dokumentix.ub.uni-siegen.de/opus/frontdoor.php?source_opus=398&la=de)

Rettberg, Jill Walker. 2012. “Electronic Literature Seen from a Distance: The Beginnings of a Field” *Dichtung Digital*, 41. <http://www.dichtung-digital.de/en/journal/archiv/?postID=278>

Rettberg, Scott. 2013. “An Emerging Canon? A Preliminary Analysis of All References to Creative Works in Critical Writing Documented in the ELMCIP Electronic Literature Knowledge Base.” Presented at *ELO2013: Chercher le texte*, Paris August 24-27.

Walker, Jill. 2002. "Links and power: the political economy of linking on the Web" *Proceedings of ACM Hypertext '02*. Pages 72-73. <http://dl.acm.org/citation.cfm?id=513358>

Graham, Shawn. 2013. "Converting 2 mode Networks with Multimodal Plugin for Gephi." <http://electricarchaeology.ca/2012/04/04/converting-2-mode-with-multimodal-plugin-for-gephi/>

## List of Dissertations Analysed

This paper discusses visualisations of references to creative works in the following dissertations. We have information about over thirty more that will be included in future versions of this paper, which of course may change the results. This paper should thus be seen as a first stage of the research. A few dissertations were not included because we lacked access to the full text or were not able to understand the language.

Mette-Marie Zacher Sørensen	Digital Poesi. Æstetisk Analyse og det Mediales Rolle i Kunstværkers Kommunikation	2013
Daniele Giampa	Oltre i confini del libro. La letteratura italiana nell'era digitale	2012
Holly Dupej	Next Generation Literary Machines: The "Dynamic Network Aesthetic" of Contemporary Poetry Generators	2012
Jeneen Naji	Poetic Machines: an investigation into the impact of the characteristics of the digital apparatus on poetic expression	2012
Jennifer Roudabush	Theorizing Digital Narrative: Beginnings, Endings, and Authorship	2012
Ugo Panzani	"I think, therefore I connect". Database, connessionismo ed esopoiesi nel romanzo anglo-americano (1995-2011)	2012
Anders Sundnes Løvlie	Textopia: Experiments with Locative Literature	2011
David Jhave Johnston	Aesthetic Animism: Digital Poetry as Ontological Probe	2011
Giovanna Di Rosario	Electronic Poetry: Understanding Poetry in the Digital Environment	2011
Jukka Tyrkkö	Fuzzy Coherence: Making Sense of Continuity in Hypertext Narratives	2011
Luciana Gattass	Digital Literature: Theoretical and Aesthetic Reflections	2011
Rulon Matley Wood	Hypertext and Ethnographic Representation: A Case Study	2011
Talan Memmott	Digital Rhetoric and Poetics: Signifying Strategies in Electronic Literature	2011
Leonardo Flores	Typing the Dancing Signifier: Jim Andrews' (Vis)Poetics	2010
Van Leavenworth	The Gothic in Contemporary Interactive Fictions	2010
Andrew Hutchison	Techno-historical Limits of the Interface: The Performance of Interactive Narrative Experiences	2009
Daniel C. Howe	Creativity Support for Computational Literature	2009
Florian Hartling	Der digitale Autor. Autorschaft im Zeitalter des Internets	2009
Wilton Azevedo	Interpoesia: o Início da Escritura Expandida	2009
Zuzana Husárová	Písanie v interaktívnych médiách. Digitálna fikcia /Writing in the Interactive Media. Digital Fiction	2009
Cheri Crenshaw	Exploiting Kairos in Electronic Literature: A Rhetorical Analysis	2008
D. Fox Harrell	Theory and technology for computational narrative: an approach to generative and interactive narrative with bases in algebraic semiotics and cognitive linguistics	2007

Jeremy Douglass	Command Lines: Aesthetics and Technique in Interactive Fiction and New Media	2007
Maria Engberg	Born Digital: Writing Poetry in the Age of New Media	2007
Nick Montfort	Generating Narrative Variation in Interactive Fiction	2007
Noah Wardrip-Fruin	Expressive Processing: On Process-Intensive Literature and Digital Media	2006
Edward Maloney	Footnotes in Fiction: A Rhetorical Approach	2005
Anna Gunder	Hyperworks: On Digital Literature and Computer Games	2004
Jill Walker Rettberg	Fiction and Interaction: How Clicking a Mouse Can Make You Part of a Fictional World	2003