Visualizing Electronic Literature Collections

Urszula Anna Pawlicka
University of Warmińsko-Mazurski

Follow this and additional works at: http://docs.lib.purdue.edu/clcweb

Part of the Comparative Literature Commons, Digital Humanities Commons, and the Other Film and Media Studies Commons

Dedicated to the dissemination of scholarly and professional information, Purdue University Press selects, develops, and distributes quality resources in several key subject areas for which its parent university is famous, including business, technology, health, veterinary medicine, and other selected disciplines in the humanities and sciences.

CLCWeb: Comparative Literature and Culture, the peer-reviewed, full-text, and open-access learned journal in the humanities and social sciences, publishes new scholarship following tenets of the discipline of comparative literature and the field of cultural studies designated as "comparative cultural studies." Publications in the journal are indexed in the Annual Bibliography of English Language and Literature (Chadwyck-Healey), the Arts and Humanities Citation Index (Thomson Reuters ISI), the Humanities Index (Wilson), Humanities International Complete (EBSCO), the International Bibliography of the Modern Language Association of America, and Scopus (Elsevier). The journal is affiliated with the Purdue University Press monograph series of Books in Comparative Cultural Studies. Contact: <clcweb@purdue.edu>

Recommended Citation
Pawlicka, Urszula Anna "Visualizing Electronic Literature Collections." CLCWeb: Comparative Literature and Culture 18.1 (2016):
<https://doi.org/10.7771/1481-4374.2902>

This text has been double-blind peer reviewed by 2+1 experts in the field.

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

This is an Open Access journal. This means that it uses a funding model that does not charge readers or their institutions for access. Readers may freely read, download, copy, distribute, print, search, or link to the full texts of articles. This journal is covered under the CC BY-NC-ND license.
Abstract: In her article "Visualizing Electronic Literature Collections" Urszula Pawlicka discusses the development of electronic literature by visualizing material available in the Electronic Literature Collection <http://collection.eliterature.org/>. Her visualization of electronic literature presents a timeline with tag clouds of keywords related to works classified chronologically by dates of publication. Pawlicka's visualization includes also all keywords of the Collection (two date there exist three Collections) separately without division in the publication dates of works. Pawlicka argues that keywords turn out to be important data to demonstrate changes occurring in the history of electronic literature. Further, in her visualization of electronic literature Pawlicka discusses the three waves of electronic literature including two transitions between them.
Visualizing Electronic Literature Collections

Volume 3 of the Electronic Literature Collection (<http://collection.eliterature.org/>) (ELC3) challenges researchers to reconsider electronic literature, its history, and current condition. Increasing numbers of twitterbot, generative work, poetry, and programming languages are a symptom of changes in electronic literature, shifting from reading to production, from close reading to hyper/machine reading, from interpretation to experience, from cyberstructuralism (see Kinder) to process-oriented studies. It seems obvious that ELC3 does not belong to the second generation of electronic literature. In the discussion about electronic literature, the prefix "post" has become noticeable, implying the emergence of period "after" the second generation indicating the "end" of forms and platforms distinguishing the second wave of electronic literature (see Cramer; Strehovec; Sample).

Much has been said about the end of electronic literature, but not necessarily about its beginnings. In 1992, Robert Coover published an article under titled "The End of Books" in The New York Times (<https://www.nytimes.com/books/98/09/27/specials/coover-end.html>) and in 1993 an article titled "Hyperfiction: Novels for the Computer" (<https://www.nytimes.com/books/98/09/27/specials/coover-hyperfiction.html>) where he presented new forms of literature. However, Coover's fascination with literary hypertext was limited to the "Storrspace School" and to "postmodernist" hypertext on floppy disks and CD-s in the Eastgate System. Thus it is no surprise that he proclaimed the "passing of the golden age" of electronic literature in 1999 ("Literary Hypertext" <http://nickm.com/vox/golden_age.html>). To Coover, the advent of internet was associated with the end of literary hypertext while for others it was the time of entering into the second generation of electronic literature (see, e.g., Hayles; Kirschenbaum; Pressman; Ryan; Wardrip-Fruin). Theorists pointed to two generations of electronic literature distinguishing by the appearance of the world wide web in 1994. This date does not raise any doubts. Now, it seems indisputable that electronic literature before 1994 is different than after 1994. With moving to the world wide web, electronic literature developed its platforms and forms: html, flash, shockwave, animation, kinet-ic poetry, and audio to name just a few. The emergence of world wide web turned out to be an event setting the boundary after which electronic literature developed differently.

Determining the next date or event indicating the third wave of electronic literature turns out to be a contentious issue. It is hard to set an exact date in analogy to the appearance of the world wide web in 1994. The emergence of the third wave is a consequence of many changes occurring between 2007 and 2009, a time of the advancement of new technologies, new mobile media, social media, and new fields (software studies, platform studies, code studies, digital humanities). All of these events contributed to the dynamic development of electronic literature that expands not only technological and artistic capabilities, but also geographical borders (i.e., the phenomenon of globalization of electronic literature). I understand the period as the transition into next waves of electronic literature: 1995-1997 and 2007-2009. My periodization is in a certain sense subjective owing to the fact that I work with qualitative texts instead of quantitative data (see Pawlicka <http://dx.doi.org/10.7771/1481-4374.2619>). However, the objective of my study at hand is to show that the assumption about waves of electronic literature is not based on to me questionable theoretical considerations, but argumentation. To produce reliable knowledge in the humanities, it seems necessary to convert text into data. Based on digital humanities practice, I investigate electronic literature as "data" and then arrive at theoretical considerations through practice.

To corroborate the idea about three transitions in the history of electronic literature, I refer to digital humanities that offers a number of tools to "produce new knowledge through the aesthetic provocation" (Jefferson 168). Following Lev Manovich's work, the method of visualization is not just a communicative act, but also an interpretative one enabling to identify, clarify, and understand data ("What is Visualization?" <http://manovich.net/index.php/projects/what-is-visualization>). In my opinion, data visualization is a good method to represent theoretical concepts in a graphical way. Taking a form of written text converted into quantitative data depicts in a visual way where icons, rectangles, fonts, and colors indicate objects and relations between them. Data visualization is a tool to exhibit the history of electronic literature and I posit that this relates to the "distant reading" of history rather than "close reading" of particular works. Therefore, I apply Franco Moretti's method of "distant reading" to look at electronic literature as an entire object and argue that this way we can grasp differences and make a division in particular periods. The first step to visualizing the history of electronic literature is to convert materials into data. To put differently, we need to know what kind of data we have, what kind of texts can be computed, and what kind of information we want to exhibit and hence the first process includes the selection, compilation, and analysis of data. The goal of the second step is to visualize, manipulate, and design data and the third step is to interpret the visualization or at least present a guide to following and reading the graphs. Thus, data visualization is a complex process, including methodological considerations, design, and theoretical approaches.
The material I use in order to discuss the development of electronic literature can be found in Electronic Literature Collection <http://collection.eliterature.org/> released in 2006, 2011, and 2016 (on the history of the collection see Pablo and Goicoechea <http://dx.doi.org/10.7771/1481-4374.2558>). I also include aspects and data of changes which occurred over last forty years in media platforms, tools, genres, and poetics. In order to grasp the shifts, I propose to investigate keywords included in each collection. Keywords help to describe, categorize, and tag digital works. We can observe the changes in the history of electronic literature by comparing keywords from each of these three collections. Consequently, keywords are data which can be organized, grouped, and visualized in the form of tag cloud to investigate and demonstrate the frequency distribution of words in each collection. However, it is not enough to just collect and count keywords: it is necessary to analyze the number of words and the relationships between them to prevent erroneous results. In this respect I take into consideration the differences between categories such as "Authors from outside North America" (only in the first collection), "Multilingual or Non-English" (in the first and second collection), and "Country" and "Language" (sections included in the third collection). While the first two anthologies were dominated by American and English works, the third publication comprises projects coming from all over the world and producing in various languages. Therefore, the solution could be to create tags such as "English" and "non-English". However, a big number of the tag "English" produces deceptive results since the rest of tags in word cloud become invisible due to their low frequency. Nevertheless, the inclusion of the category indicates the growing number of "non-English" works and this is crucial to show the process of globalization of electronic literature. I also include only the category of "non-English" understood as a work produced in a language other than English in the last part of visualization that compares keywords from three collections separately. Only by unifying such determinants is it possible to capture significant language differences between the three collections of electronic literature. The next modification regards keywords of the third collection. Visualization covers also terms used to describe digital work, but not necessarily included in the section "Keyword": it concerns terms associated with software, tools, and programming languages which present fully the third wave of electronic literature.

The following step to visualize data is the selection of program whereby we can create tag clouds. The internet offers many free online word cloud generators; however, all of them have limited features and visualization available in Wolfram Mathematica <https://www.wolfram.com/mathematica/> provides functionality to create variety of word clouds. Thanks to this program, it is possible to include two or even three of the same words with different numerical values in one graph, visualize together three groups of words differing by font color, and choose the background color. A whole graphical representation covering periodization on a timeline with particular tag clouds, is created in Prezi which enables you to easily explore and zoom any part of visualization.

Figure 1: "The Visualization of Keywords from Electronic Literature Collections vol. 1-3" (<http://www.urszulapawlicka.com>)

The visualization in Figure 1 presents three waves of electronic literature, determined by the frequency of keywords from the three Electronic Literature Collection-s (2006, 2011, 2016). Works from each collection are divided into five groups based on the year of their publication. My visualization shows three waves of electronic literature with two transitions between them: the period until 1995...
The first visual representation of information relates to the publication dates of works included in each collection. By following graphic lines, we can notice from which years works are available, which years can be hailed as a "boom" of electronic literature, and when it has started the turn toward the restoration of old projects. The pink line indicates works from the first collection which covers one project published before the year 1995 and eleven projects released between 1996-1999. Projects originate from the years 2000-2006 with the peak in the years 2004-2005 (twenty-two works). The turquoise line marks the second collection where only one work comes from years before 1995. The next project dates from 1998 and the rest of works come from the years 2000-2010 with the boom in 2008 (twenty projects). The green line signals works included in the third collection. At this point, it is noticeable that changes related to the shift toward the reconstruction of electronic literature originated before the advent of the world wide web and hence the third collection covers eight projects produced between 1975-1993. Electronic Literature Collection includes works published up to 2016 with a peak in the years of 2012-2014 (sixty-three projects). Analyzing this part of visualization, we need to be aware that collections do not include hypertexts published in Eastgate System in the 1990s. We have not still had access to many digital works released on floppy disks or CD. However, the visualization supports the argument regarding the turn toward archiving and restoration of electronic literature. It is interesting to look at the boom years of electronic literature occurring nearly every four years. Each of these peaks is in a certain sense an effect of the development of technology and media platforms: the domination of Flash in 2004-2005, locative works, 3D, Java in 2008, and Javascript, Netprov, Twitter, Twine, Ruby, and Python prevailed in the years 2012-2014.

The next part of visualization is devoted to the presentation of tag clouds listed on the timeline. First, I investigate the frequency of keywords without division and then with divisions in the collections. The application of general characteristics can address the issues such as shifts in the list of keywords occurring within forty years (deleting, modifying, and adding words), the changes in the development of electronic literature genres (dominant particular genres in each wave), the transitions in digital tools and media platforms, and the relationships between media platforms and literary forms. The visualizations with the division into three collections, in turn, can provide information about the development of new media poetic and theoretical language of electronic literature, the process of filling the gaps in the history of electronic literature, the process of restoration of electronic literature that broadens the knowledge about the beginnings of digital literature, as well rewriting the history of electronic literature where hypertexts in the 1990s are not its first forms.
The first wave of electronic literature lasted up to 1995 (see Figure 2) and it is the time of literary experiments seeking to create automatic and generative poetry as a continuation of the Oulipian conception of combinatorial literature. Tag cloud includes the keywords related to software (BASIC, HyperCard), literary forms and genres (codework, poetry, generative, hypertext, combinatorial, performance/performative, animation/kinetic, visual poetry or narrative), and theoretical categories (non-interactive, critical/philosophical/political, network forms, text movie, retro, Oulipo, database, activist). Due to the fact that I have no access to many works from said period, the graph consists of a small number of keywords. However, tag cloud supports previous arguments regarding the domination of hypertext (because of the lack of access to some hypertext from "Storyspace School," word "hypertext" does not reflect its significant role at that time) and codework, as well as the function of programs existing before the advent of the world wide web.

Tag cloud with the division into three collections shows that actually the third collection has led to extending the knowledge and categories regarding the first wave of electronic literature. So far, just only two pieces of works provided limited information about that period, including codework, critical/philosophical/political, network forms, non-interactive, performance/performative (pink font tags), and animation/kinetic, poetry, visual poetry or narrative, non-interactive, retro, text movie (blue font tags). The third collection expands this list of words by publishing 8 projects tagged by the following categories, such as poetry, code, activist, generative, combinatorial, BASIC, Oulipo, hypertext, database, and HyperCard.

Along with the movement to the world wide web, electronic literature has developed its new digital tools, media platforms, and forms (see Figure 3). The time interval between 1995-1997 is called as a transition to the second wave of electronic literature when we can observe significant shifts in practice and theory of electronic literature (see Aarseth). Tag clouds present the following categories (order from highest to lowest numbers) associated with programming languages and platforms (html/dhtml, Flash, Javascript, Storyspace), forms (hypertext, poetry, fiction, visual poetry and narrative, gif), and concepts (video, audio, collaboration, memoir, women authors). It is the time of the emergence of new tools (Flash, html, Javascript), and elements of electronic literature that started combining text with audio and video. This tendency led to the formation, development and domination of Flash poetry, visual poetry, animation and kinetic in the second wave of electronic literature. This picture of the transition of electronic literature is mainly shaped by the third collection. By analyzing tag cloud with the division into three collections, we can notice differences between anthologies regarding the years 1995-1997. So far, two works from that period were published in the first collection, described by the categories: hypertext, html/dhtml, fiction, visual poetry or narrative, Storyspace, audio, collaboration, memoir, and women authors. Three works, in turn, are released in the third collection, expanding keywords by adding new categories: Flash, Javascript, video, and gif. Therefore, thanks to the third collection, we can tell more about the development of particular tools in the time of transition to the second wave of electronic literature.
The subsequent part of visualization relates to the second wave of electronic literature from 1998 to 2006 (see Figure 4). It is the time of establishment of Electronic Literature Organization in 1999, the dynamic development of electronic literature genres, theories, new media poetic (a significant increase in the number of categories to describe digital works), and the growing number of web-based software (Flash, Shockwave, QuickTime) and programming languages (Javascript, Java, Processing, Perl). Tag cloud presents the following categories (listed in the order from highest to lowest numbers) related to technologies, software and programming languages (Flash 41 times, Shockwave 13, JavaScript 9, html/css 8, Java 5, QuickTime 5, Processing, Inform, Squeak, Cave, VRML, TADS, Viral, Perl, mobile, augmented reality, virtual environment, GPS, spatial augmentation); genres and forms (poetry 26 times, visual poetry or narrative 24, animation/kinetic 24, generative 20, fiction 15, games 14, combinatorial 12, hypertext 11, conceptual 7, 3D 7, interactive fiction 6, virtual environment, GPS, spatial augmentation); and categories (audio 37 times, women authors 24, collaboration 28, textual instrument 13, place 10, appropriated texts 10, critical/political/philosophical 10, parody/satire 9, network forms 8, non-fiction 7, time-based 7, memoir 6, wordtoy 6, ambient 6, narrative 6, stretchtext, text movie, translation, documentary, music, essay/creative non-fiction, children's literature, gender / race / sexuality, constraint-based/procedural, database, mash-up, video, web-based, hacktivist).

In sum, the second wave of electronic literature is distinguished by the dominance of web-based software, especially Flash whose number has significantly decreased from 41 to 15 (in 2007-2009) and 9 times after 2010. Platforms accessible by the internet led also to the development of works consisted of audio, video, music, animation, kinetic, etc. Moreover, focus on new technological possibilities contributed to the production of new poetic categories emphasizing new artistic capabilities such as "textual instruments," "stretchtexts," "text movies," "wordtoys," etc. My attention to the second way of electronic literature was to describe and analyze new media language and new media poetics (see Glazier; Funkhouser; Hayles; Kac; Manovich; Morris and Swiss). The third wave of electronic literature, in turn, departed from theoretical construction toward production and operation. Consequently, the frequency of categories from the second wave has drastically decreased in the next phases and some completely deleted. For instance, "audio" has dropped from 37 to 12 (in 2007-2009) and 10 times after 2010; "visual poetry or narrative" lowered from 24 to 6 (in 2007-2009) and eliminated from the third collection; "animation/kinetic" decreased from 24 to 6 (in 2007-2009) and removed in the last collection. The following categories, in turn, are deleted from the discourse of electronic literature: "stretchtext," "appropriated texts," "ambient," "text movie," "textual instrument," etc.

The first and second anthologies used nearly the same group of keywords and this means that the main differences take place in the third collection. New categories, added in the last collection, demonstrate that the goal of electronic literature collections is not only archiving current works, but also digging into the history of electronic literature to show that some digital forms are not new at all, but emerged earlier. The best example of it is the term "locative" from the second collection referring only to projects produced in the year 2007-2009. Only the third collection releases the locative work created in 2002 during the second wave of electronic literature (see Hight, Knowlton, Spellman <http://collection.eliterature.org/3/work.html?work=34-north-118-west>). The next instance relates to the keyword "gif" that emerges the first time in the third collection. Although animated gif shows up...
with the rise of mobile and other platforms, the editors of the last collection apply this term to describe works produced in 2004 (see Ni_ka <http://collection.eliterature.org/3/work.html?work=hallelujah>) and even in 1997. Typoemas created by Ana Maria Uribe in 1997 was depicted as an animated concrete poetry (<http://collection.eliterature.org/3/work.html?work=tipoemas-y-anipoemas>) and now this work is tagged as an example of gif. Taken together, by analyzing the relationships between three collections, we can perceive how the language of electronic literature has changed and how it has rewritten theories of digital literature.

Figure 5: Keywords related to works published during the years 2007-2009 (left panel). It corresponds to label 1d in Figure 1. Keywords with division into three collections (right panel). It corresponds to label 3d in Figure 1. Turquoise indicates the second collection and green the third collection (<http://www.urszulapawlicka.com>).

The time interval between 2007 and 2009 is called as a transition to the third wave of electronic literature (see Figure 5), a time of improvement in new technologies, the emergence of social media, and the establishment of new fields of artistic production. These events led to the development of new stage of electronic literature hailed as a "post-hypertext e-literature" and "post-Flash area." General word cloud includes the following categories referred to media, technologies, and programming languages (Flash 15 times, Javascript 7, Java, Shockwave, Ruby, mobile, Cave, Processing, augmented reality, virtual environment, Twitter, Bot), genres and forms (poetry 15 times, generative 9, visual poetry or narrative 8, hypertext 6, animation/kinetic 6, codework, games, 3D, ergodic, locative, combinatorial, conceptual, performance/performative, installation, remix, interactive fiction), and poetic categories (audio 12 times, collaboration 10, critical/political/philosophical, narrative 9, network forms, database, textual instrument, video, text movie, essay/creative non-fiction, documentary, parody/satire, hacktivist, non-interactive, constraint-based/procedural, appropriated texts, retro, place, memoir, gender/race/sexuality, mash-up, ambient, wordtoy, activist). For the first time, there emerge new concepts typical for the third wave of electronic literature: Twitter, Bot, remix, etc.

Works produced in 2007-2009 are published in the second (33 works) and third collection (9 works). Although most of the projects come from the second collection, the above mentioned new categories emerge only in the last collection. This observation again supports the idea that the third collection aims to "dig out" old works which shed new light on the previous years of electronic literature. The last collection allows us to look at that period not as a continuation of the second wave, but as a transition to the next phase of electronic literature by introducing works with new media platforms.
The last tag clouds on the timeline present the third wave of electronic literature that has started since 2010 (see Figure 6). The graphs covers the categories related to media, platforms and programming languages (Javascript 31 times, Twitter 13, Bot 10, html 9, Flash 9, Twine, mobile, Kinect, Java, email, virtual reality, Prezi, Cave, augmented reality, CSS, jQuery, Processing, AutoCAD, Python, Ruby, Scalar), forms (poetry 36 times, generative 32, hypertext 15, game 11, performance 6, code 6, installation 5, comics, fanfiction, interactive fiction, Netprov), and categories (remix 15 times, networked 13, audio 10, activist 7, Oulipo, web-based, database, video, hacktivist). The third wave of electronic literature is fueled by technological development and social media. Artists moved from using previous software toward new platforms and programming languages whose usage increased dramatically. For instance, the frequency of the keyword "Javascript" has jumped from 1 time (1995-1997), 9 (1998-2006), 7 (2007-2009) to 31 times (after 2010). The emergence of social media led to the expansion of electronic literary forms, including Twitterbot, fanfiction, and remix. The development of new technologies and new artistic capabilities cause that artists focus on production and the processing of works instead of text and content. A currently developing form is the artificial intelligence novel and this relates to the question of the "automatization" of literature. Artists concentrate on the creation of generators to produce works and this is noticeable in the growing number of keyword "generative" from 4 times (1975-1995), 20 (1998-2006), 9 (2007-2009) to 32 times (after 2010). Consequently, it contributes to the emergence of new form of electronic literature, namely "e-poem as a platform" where artistic work acts as a platform to generate next poems. The best example of it is Nick Monfort’s Taroko Gorge (<http://collection.eliterature.org/3/work.html?work=taroko-gorge>) which became an open-source "platform" to remix and generate work. The third wave of electronic literature is also distinguished by a tendency to narrow concepts. In the light of transmedia literature, it is hard to classify works created at the intersection of various kind of arts. Therefore, we can observe the growing number of general concept "poetry" encompassing different kind of expressions, as well as the departure from the previous categories with the obscure meaning ("stretchtext," "wordtoy," "ambient," etc.).

Most of the works created after 2010 is included in the third collection and because of this, the tag cloud with the division into collections does not differ significantly from the previous graph. Further, this one work from the second collection is called as a "codework" and this supports the image of the third wave of electronic literature.
The last part of my visualization shows tag clouds of keywords from each collection separately without division into the publication date of works (see Figure 7). The pink font indicates the first collection, blue font the second, and green the third collection. The key difference relates to the number of categories included in the section "Keyword": the first collection covers 51 terms, the second 50, the third 36. These values support the assumption regarding the reduction of theoretical concepts and departure from new media poetics. The next inference refers to the comparison of category "non-English." It is not hard to notice a gradual increase of non-English works from 1 project included in the first collection, 14 in the second collection to 48 works covered in the third collection (number of works from the "Language" section). By adding this tag, it becomes possible to show the phenomenon of globalization of electronic literature that has started to be one of the primary goals of electronic literature communities and topics in the discussion of digital literature.

Volume 3 of Electronic Literature Collection confirms the notion of transition to a "post" area. However, it is difficult to prove theoretical assumptions without empirical evidence. In a sense, the humanities is free from the proof of theorem since scholars operate with qualitative instead of quantitative texts. Nevertheless, the digital humanities has moved the scholars closer to the science through the development of tools and methods to support theoretical considerations. One of these techniques is data visualization that turns out to be an adequate method to show and convince that the advent of the third wave of electronic literature has become real.

Note: I thank Aydin Deger (Stony Brook University) for introducing Wolfram Mathematica to me to generate visualizations.

Works Cited

Ni_ka. "__Hallejujah" *collection.eliterature.org* (2004):


Author's profile: Urszula Pawlicka is working towards her doctorate in Polish philology at the University of Warmińsko-Mazurski with a dissertation entitled *Literature and New Media: Theory and Practice*. She is also member of the Laboratory of the Intersemiotic and Intermedia Research at the University of Warsaw. In addition to numerous articles, Pawlicka published the single-authored book *(Polska) poezja cybernetyczna. Konteksty i charakteryztyka* (2012) ([Polish] Cyberpoetry: Contexts and Characterization) and the edited digital volume *Cybernetic Poetry in Poland* <http://issuu.com/korporacja_haart/docs/polish_cybernetic_poetry_refresh> (2013).

E-mail: <ulapawlicka@gmail.com>