

Dynamics of Remembering and Forgetting on the (Social) Web Platforms

Simeona Petkova
University of Amsterdam, The Netherlands
dcmona@gmail.com

Abstract

In recent scholarly works, the Web is viewed either as preserving all the available data for unlimited period of time or forgetting it all in a short-term succession. “[...] because of digital technology, society’s ability to forget has become suspended, replaced by perfect memory” (Mayer-Schonberger, 2009) while at the other extreme, the digital cultural heritage is “at risk from loss of data, knowledge or memory” (Blome and Wijers, 2010). This paper proposes to examine the dichotomy by exploring several (social) Web platforms. The article takes upon how the Web’s ability to remember and / or forget is often approached through studying the Internet Archive, or the personal involuntarily collected histories (and their relation to privacy issues, search engine back-end politics, identity, markets, users generated content, etc.), or the ways the Web is changing the way we think and remember. It proposes to focus instead on the Web (social) platforms by asking what is the medium-specific way to study what is preserved or left behind? The article proceeds with analysis of two projects carried out within the digital methods framework and maps the contribution in the ways the medium-specific analysis explicates the dynamics of both remembering and forgetting through revealing mechanisms that reconfigure and reshape memory as content (memory narratives).

Keywords

remembering, forgetting, Web platforms, medium-specific analysis, digital methods, reworking

1. Introduction

Digital remembering and forgetting have a potential to “be crucial for knowledge and power distribution in the future” (Anna Maj & Daniel Riha (2009:2) since our physical journals, letters, and photographs have been substituted by “fresh masses of life-affirming digital stuff: five billion images and counting on Flickr; hundreds of thousands of YouTube videos uploaded every day; oceans of content from 20 million bloggers and 500 million Facebook members; two billion tweets

a month”¹. Instead of focusing on the limitations of the web archiving and the potential lost of online content / knowledge / cultural heritage, this article asks how is the dynamic of remembering and forgetting negotiated on the Web’s platforms and how do platforms rework / reshape past memory?

The dynamics of remembering and forgetting is normally studied through the Web’s archives and the limitations they impose on the preservation of the online content; through the abundance of traces and memory of the users (and the implications for privacy); and through the impact of the Web on what and the way we remember. This article proposes a medium specific approach for studying the dynamics of remembering and forgetting that focuses on the ways the Web platforms rework and retell past memory. Two case studies are presented and discussed here: “Historical Controversies Now” is an empirical study on several historical controversies and the ways the Web’s dominant platforms rework them. “Neutral or National Point of View? Wikipedia’s language versions as cultural reference”, undertaken by Emina Sendijarevic (MA student of “New Media & Digital Culture” program, University of Amsterdam), traces the ‘deviations’ of the Wikipedia entry of ‘Srebrenica’ across language Wikipedias.

The analysis found that different platforms have specific dynamic of remembering. Twitter and Google News actively rework and make present various historical events. Google Search falls in the category of retrieving sources and documents (Wikipedia articles, history sites) without reworking them. The social Web platforms (YouTube, Flickr, Facebook) actualize the past depending on the event itself. However, Flickr sources were found to contest the events more than YouTube which is a pattern common across the histories. Another finding is attributed to the mechanisms by which the reworkings are undergone. Especially in the case of Wikipedia, they are exclusively enabled by the platform itself: specific language entries, citations for verification, recognizable framing in the cases of controversy, anonymous edits and forking.

The article unfolds as the following: Section 2 overviews the related study of Web archives, digital traces / metadata, and the impact of the Web on the ways we remember. Section 3 introduces the methodology followed by the presentation and the discussion of the two case studies. Finally, section 5 maps the findings and the contributions.

¹ R. Walker (2010). “Cyberspace When You’re Dead.” *New York Times Magazine*, 01: 2010, http://www.nytimes.com/2011/01/09/magazine/09Immortality-t.html?_r=4&pagewanted=all, retrieved 12 January 2010.

2. Related Work

2.1 Web Archives

A common approach for studying the dynamics of Web's ability to remember and / or forget is to focus on the online archives and their repositories of time-stamped versions of webpages taken over a long temporal period. The Internet Archive, for example, provides snapshots of the sites-in-time (url based) through the Wayback Machine. Rogers (2010) and Weltevrede (2011) distinguish three types of historiographies enabled through the online archives: single-site histories / biographies of the websites (via the Wayback Machine); event-based histories (via the Library of Congress special collection); national histories (via the national libraries Web collections). Capturing the webpages histories in time (sites-centric approach) does pose significant questions concerning the reliability of the archiving process and the amount of Web's data that has been archived or left behind.

In "Website history and the Website as an object of study" Brügger defines the website not as an entity experienced by a concrete visitor, but as an object of analysis in its own, that's in embedded in its media environment, its textual environment, and its textuality. He makes a detailed and descriptive step-by-step methodology of studying the website acknowledging its relation to the Web in general (through hyperlinks), to the browser in which it is viewed and the morphological, syntactical, and semantic interactions between its sub-elements ("...one can distinguish the following five analytical strata: the web as a whole; the web sphere; the individual website; the individual webpage; and an individual textual web element on a webpage, such as an image"²). Brügger acknowledges that the process of archiving is always contested with the dynamic of updating resulting in either uncertainty whether all the elements have been archived or danger of preserving something that was not initially there. This invites subjective reconstruction on the researcher's side since it employs technical and temporal deficiencies (missing words, images, graphics, dynamics of updating, etc.). Schneider and Foot (2004) also propose an integration of the scholar into the archival process: "The emergence of web archiving techniques that are designed to facilitate scholarly analysis integrates researchers into archiving activities"³

Following the Wayback Machine's specificity to archive and retrieve the websites through their host url, Michael Stevenson⁴ also conceptualizes several archival limitations. The navigation in the Internet Archive 'jump-cuts' through time due to issues connected not only to the limited

2 N. Brügger (2009). "Website history and the Website as an object of study." *New Media & Society* 11.1-2: 115, pp. 290

3 S. Schneider and K. Foot (2004). "The web as an object of study." *New Media and Society* 6.1 (2004): 114-122, pp.119

4 M. Stevenson (2009), "The Archived Blogosphere, 1999-2001" in *Changing Cultures: Cultures of Change*, draft conference paper

indexing and hard drive space, but to blocked sites, outdated software and formats that reject the archiving process. If there is a missing snapshot of the site-in-time, the Wayback machine maintains the continuity of access by connecting to the url in the present. This leads Stevenson to conclude that the Internet Archive as a legacy system still employs the major principles of the cyberspace – navigation through urls and editorial approach of building archival collections.

Concerning the amount of Web's data that has been archived or left behind one can ask how (by following the archival medium) to historically research real-time platforms such as Twitter or Facebook? Since the Web archives contain time stamped versions of the web page over a period of time, can these sites be reduced to a single history of their front pages (log-ins)? Wendy Chun (2008, 2010) explicitly criticizes (in a more generic way) the ability of the Web and the digital media to preserve and archive: digital media is erasable, forgetful, sources disappear. "The belief in the Internet as a cultural memory, paradoxically, threatens to spread the lack of memory everywhere and plunge us negatively into a way way back machine: the so called digital dark age" (2008: 169). Digital cultural heritage is "at risk from loss of data, knowledge or memory" (Blome and Wijers 2010) and Chun (2010) proposes undertaking of activist practices of continuous preservation of our cultural memories due to the 'ephemeral nature' of the Web.

2.2 The Web Remembers It All

Another line of research related to the Web's ability to remember and / or forget considers the online as preserving all the data for unlimited period of time "[...] because of digital technology, society's ability to forget has become suspended, replaced by perfect memory" (Mayer-Schonberger 2009). The dominant operational mode is 'to preserve' since processing and storing of information becomes easier and inexpensive. Companies like Google collect and retain information about their users' search behavior and their digital traces for years. To balance what could be remembered and what forgotten (and for how long), Mayer-Schonberger proposes an expiration date for digital information to be incorporated. This line of research comes close to the notion of the 'data body' coined by Critical Art Ensemble (CAE). The 'data body' is the overall collection of files connected to an individual that has two major functions: to serve a controlling (repressive) or a marketing apparatus. It provides accurate demographic information to marketers for creating target audiences.

The data body gives them insights into consumption patterns, spending power, and "lifestyle choices" of those with surplus income. The data body helps marketers find you, and provide for your lifestyle⁵.

⁵ Critical Art Ensemble (1995). "Utopian Promises – Net Realities," (Hamburg: conference proceeding of *Interface 3*), pp. 146.

Another criticism connected to the retention politics of the Web's dominant devices is the search engines critique of their ability and the practice of tracking and preserving online activity for years. During his presentation at the 2009 edition of "Impakt Online Festival" Richard Rogers argued that in the different stages of the Web (from hyperspace to cyberspace to social space) a clear tendency of the 'demise' of the browser could be traced. The history of the query has migrated from the desktop browser to the search engine cloud. By preserving the search histories, Google profiles or creates data bodies of its users. In "Google's Personalized Results: the 'New Normal' that Deserves Extraordinary Attention"⁶ Danny Sullivan elaborates over Google's introduction of the 'personalized search' in December 2009. By saving the users' queries for one hundred and eighty days, Google customizes the search results in accordance to them. Previously, the personalized search was enacted when the user is logged in a Google account while performing the search. After the innovation, everyone who has not opted out the personalization search receives the Google results tailored in accordance to their queries history. There are several aesthetic responses to this power related phenomenon. *Scroogle* is a browser plug-in that prevents Google from setting cookies, seeing the IP address of the user and registering his/her search histories. *I Love Alaska* (Figure 5) is a series of mini-movies that depict the AOL data leak through the queries of user 711391. In 2004, an AOL text file containing twenty-million search keywords for over 650,000 users collected for a three-month period appeared on one of its websites and circulated the Web. Each user was not identifiable through his/her personal information but was listed under a unique sequence that enabled a researcher to compile the user's search history.

For Esther Weltevrede (2011) the two dominant discourses (of the Web that is in danger of losing content and the Web that has abundance of digital traces) exist side-by-side since they are based on different understanding of data. The first one focuses on the Web's content and its potential as a source of knowledge (challenged by the inefficiencies in the archival process) while the second one understands data as personal information / seen as metadata. The article continues with presenting a third line of research related to the ways dynamics of remembering and forgetting on the Web is approached and studied.

6 D. Sullivan (2009). "Google's Personalized Results: The 'New Normal' that Deserves Extraordinary Attention" (published online <http://searchengineland.com/google-personalized-results-the-new-normal-31290>, 2009), retrieved 8 February 2010.

2.3 Mind the Web

In “The Shallows: What the Internet is Doing to Our Brains” Nicolas Carr undergoes a historical analysis to trace the impact of several tools we use to find, store, and share information on the ways we think and remember. He examines the map, the clock, the book, and the Web. For the purposes of the current argument I will focus on the Web and its implication on the ways we remember.

Briefly summarizing, Carr argues that our brains adapt to the ways we retrieve information on the Web. The tools we use to find, store, analyze and share information are seen as an interruption system based on a fast pace of receiving and generating responses. According to Carr, on the Web we are constantly distracted by hyperlinks, updates, and messages that translates into a change of our habits and in our attention span: we ‘skim’ pages instead of attentively reading them. The Web’s informational abundance strengthens the visual-spatial intelligence, our abilities to gather and filter information, to express and collaborate but weakens the knowledge acquisition (people who read a linear text comprehend, remember, and learn more than those reading hyperlinked text), the critical thinking and reflection.

Moreover, according to Carr our working memory is in a permanent state of overload. Our knowledge relies upon the ability to transfer information from the working to the long-term memory where the impressions, ideas, and experiences gain depth and complexity. However, while there is no capacity for our long-term memory, the working one is limited: we are able to hold some amount of information at a time. The small capacity of the working memory is what makes our short-term retention vulnerable to distractions. For Carr a break in our attention could result in inability to transfer information to the long-term memory, i.e in forgetting. The small capacity of the working memory also makes it prone to overload: when the information exceeds the storage, we become unable to relate and translate it to our lasting memories. The mind loses its ability to process and store what we perceive and this weakens the ability to understand, learn and compile knowledge. Carr presents a case study to support the above: in 2009 researchers at Stanford carried out cognitive tests on two groups of Web users: ‘heavy’ and ‘light’ multitaskers. They have found that the users prone to more and dynamic Web exposure tend to be more easily distracted, show lower ability to concentrate on a task and have less control over their working memory in comparison to the ‘light’ online multitaskers.

Carr’s approach relies on cognitive sciences data and case studies to make its claims on the ways the Web reflect our thinking and ability to remember. In the same school of thought is Gabriela Taddeo for whom the social media platforms (together with media and cultural

consumption) strongly reflect “the cognitive, cultural and affective moods with which people remember, choose what to remember and/or share, and preserve and socialize memory.”⁷ Their findings are important for studying the dynamic relations between the Web and memory although they are primarily focused on the effects of the medium not on the medium itself.

After discussing the common approaches to studying the dynamics of online remembering and forgetting (the focus on the Web’s content and the limitations of the online archives; the abundance of digital memory / traces and their implication for privacy; the effects of the Web on the way we remember), the article proceeds with introducing a new approach to study this dynamic. It builds upon the focus on the content but it does not deal with the limitations of the Web archiving. Instead, it analyzes the ability of the Web platforms to actualize and rework memory using digital methods that acknowledge the specificity of the Web as a medium. Introducing a new approach to study the online remembering / forgetting dynamic is informed by several factors. First, it reflects the recent call in media studies to move away from the solely representational analysis focused on the effects of the medium (“the mainstay of media and cultural studies approach to communication” (Parikka & Sampson eds: 2009). Second, the approach aims to contribute to the aspirations “to explore the co-relation of cultural and experiential practices, thought and intelligent devices.”⁸ and the need in memory studies to research the ways memories are mobilized and networked in relation to the Web’s (globytal) characteristics of multi-modalities and data flows that “[...] traverse the organic and inorganic, the human and the machine, the public and the private, the individual and the institutional” (Reading 2009:36).

Before presenting the two case studies that explicate this approach, I would like to briefly introduce the methodology used in the projects, i.e. the digital methods developed by prof. Richard Rogers and ‘Digital Methods Initiative’, University of Amsterdam.

7 G. Taddeo (2009). “The Blended Boundary between the Individual and Social Memory in the 2.0 Web Era,” *Observatorio (OBC*) Journal* 11: 115-131.

8 M. Fuller, “A Billion Gadget Minds: Thinking Widgets, Data, and Workflow”, email correspondence 14 June 2010.

3. Methodology

3.1 Digital Methods and the specificity of the Web as a medium

In “The End of the Virtual: Digital Methods” (Inaugural speech, 2009), Richard Rogers specifically focuses on the need for new methods to research web media. He grounds his claims through two propositions: Internet is no longer considered to be virtual realm existing independently from our social and cultural context. This reflects the questions of study: they move from “*just* online culture” to “how to diagnose cultural change and societal conditions with the Internet” (Rogers 2009: 3). Secondly, he differentiates between two types of methods that can be used to study the Web: ‘digitized’ and ‘natively digital’. Although there is no ontological difference between digitized and digital, Rogers argues that since the objects of research exist only on Internet (hyperlinks, tags, archived Websites, etc.), we need to develop ‘online’ methods that follow the ‘online’ objects. He gives examples of how this can be carried out: commonly the ‘hyperlink’ is studied through hyper textual literary theory or social networking theory with methods originally conceived outside of the digital realm. Rogers, proposes instead, the hyperlink to be studied through *Issue Crawler*, software developed specifically to evaluate inlinks and outlinks as markers of impact and reputation. Websites, on the other hand, are commonly analyzed through usability tests with registration approach. Rogers argues that they can be also analyzed as an archival object made retrievable through the Wayback Machine. For example, the history of Google’s Web directory can be captured as a line of moving snapshots explicating societal change: the replacement of the human editor by the back-end algorithm. Digital Methods Initiative has developed “online grounded” methods to study also the search engines (with *Issue Dramaturg*) or the social networking sites (as providing post-demographic data) or the Wikipedia edits (through *Wikiscanner*).

The article continues with the presentation and discussion of two case studies “Historical Controversies Now” and “Neutral or National Point of View? Wikipedia’s language versions as cultural reference” that represent a medium specific approach of studying the dynamics of remembering and forgetting on the Web’s platforms and have incorporated digital methods for their data collection.

4. Case Studies

There are two case studies presented and discussed in this chapter. “Historical Controversies Now” is an empirical study on several historical controversies and the ways the Web’s dominant platforms rework them carried out during the ‘Digital Methods Summer School 2010’⁹. The second project “Neutral or National Point of View? Wikipedia’s language versions as cultural reference”, undertaken by Emina Sendijarevic, traces the reworkings of the Wikipedia entry of ‘Srebrenica’ across language Wikipedias.

4.1 ‘Historical Controversies Now’

“Historical Controversies Now” focuses on debated historical moments charged emotionally and still being an object of contestation. The histories were related to the online time perceived as “multiplicity of times derived from relations between different elements” (Leong et.al 2009, 1279) by asking: Do we primarily find contemporary sources or historical sources in the various spheres? Does this vary across controversies? Do the sources on a platform focus on the historical moment itself, or a contemporary reworking of the moment? Does this vary across controversies?

The Wikipedia controversy page was used as the baseline for the selection of the historical controversies. Five past controversies before the year 2000 and five after 2000 were chosen with differentiation in type (revolution, bombing, disaster, etc.) and location (Europe, US, Asia). The selected historical events were as follows: French Revolution (1789); Armenian Genocide (1915); Atomic bombings of Hiroshima and Nagasaki (1945); Bloody Sunday (1972); Tiananmen Square protests (1989); 9/11 (2001); Sars (2003); Muhammad cartoons (2005-2006); Lehman Brothers bankruptcy (2008); BP oil spill (2010). The research team has also selected several dominating platforms representative for the Web’s spheres and our daily online activities: Google search (Web sphere); Google Blogsearch (Blogosphere); Search.Twitter.com (Twitter); Googlenews (News sphere); Googlebooks (Books); Google Scholar (Science); Facebook (Social networks); You Tube (Videos); Flickr (Photos).

The sample for the data collection consists of the first ten results returned by each platform per controversy (i.e. approximately 1000 results). The ranking, the date source, the title source, the description source, and the urls source of each result were recorded and coded as following: 1) code past (the result focuses on the historical event); 2) code present (the result reworks the past event¹⁰);

⁹ “Historical Controversies Now” has been carried out by Demet Dagdelen, Esther Weltevrede, Marije Rooze, Martin Feuz and Thomas Poell.

¹⁰ The source actualizes the past event by evoking / interpreting it or by problematizing it; the date of the source plays a secondary role in the coding since the time registration online is a complex and often misleading process.

3) code cold (the controversy is settled); 4) code hot (the controversy is continuously actualized). The research team underlined that the coding of the sources in ‘present’ and ‘hot’ was undertaken only when the results explicitly rework, problematize, politicize or link to a contemporary event. In some cases, the actual sources have been also reviewed when their title or the description were insufficient to determine whether the result is code present / past / cold or hot.

I will proceed with discussing the data collected for four of the past events highlighted as the most representative from the ‘Historical Controversies Now’ research team. The Armenian Genocide (Fig. 1) is found to be dynamically remembered and reworked in the present especially on Twitter, Google Blogs, and Google News platforms. Facebook also actualizes the event while for Google Books and Google Scholar this is a past and settled controversy (i.e forgetting is implied). Both Flickr and YouTube return past results with Flickr more reworking and problematizing the events. French Revolution (Fig. 2), on the other side, is considerably less actualized in comparison to the Armenian Genocide. The reworking of the French Revolution in the present takes place

Historical controversies Now

Querying historical controversies in dominant devices and platforms, the question we ask is what kind of history are we accessing on each device?
[Read more](#)

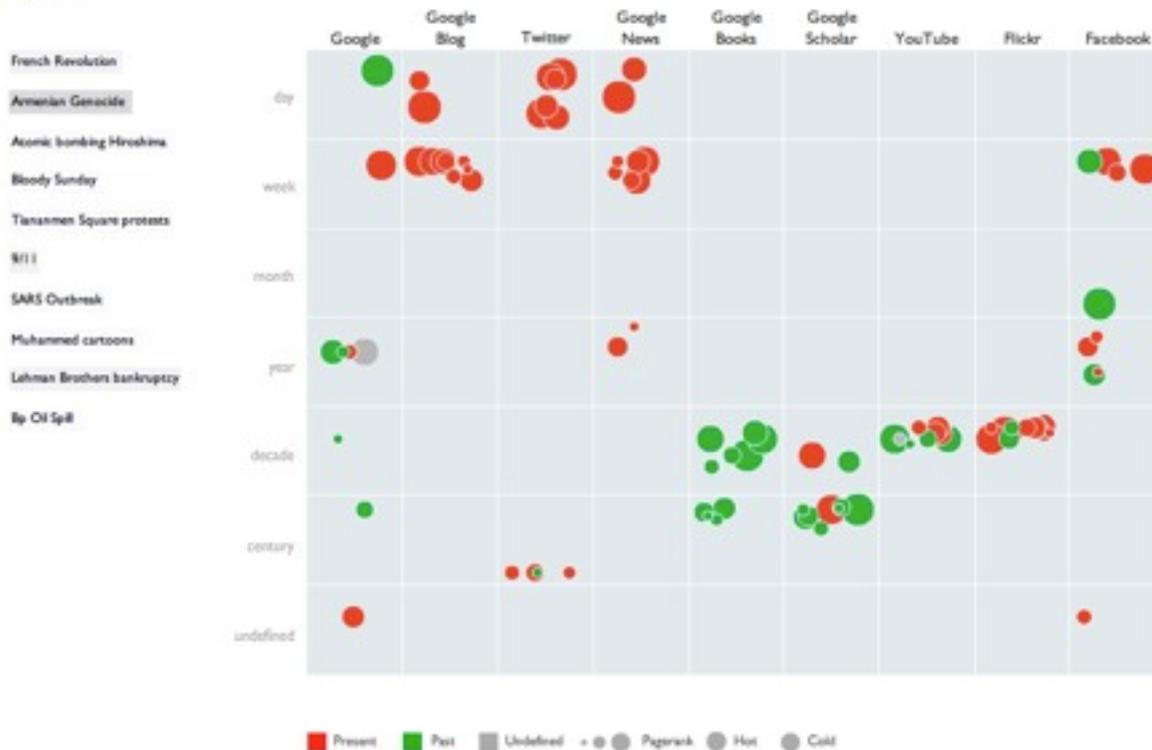


Fig. 1 Reworking of the Armenian Genocide across Web platforms

mostly on Twitter while Google, Google blogs, Google Books / Scholar and YouTube provide settled account and past sources. It is interesting to note that Flickr (again) problematizes the event in comparison to YouTube where the historical controversy is not disputed. 9/11 (Fig.3) is found to be the most dynamically actualized historical event on Twitter, Google Blog and Google News. Facebook also contributes to its present actualization and reworking while Google's sources are predominately from the past and settled. YouTube's and Flickr's reworking dynamic is with sources from a decade ago which contest the events in comparison to Google Books platform that does not dispute or problematize 9/11.

Historical controversies Now

Querying historical controversies in dominant devices and platforms, the question we ask is what kind of history are we accessing on each device?
[Read more](#)

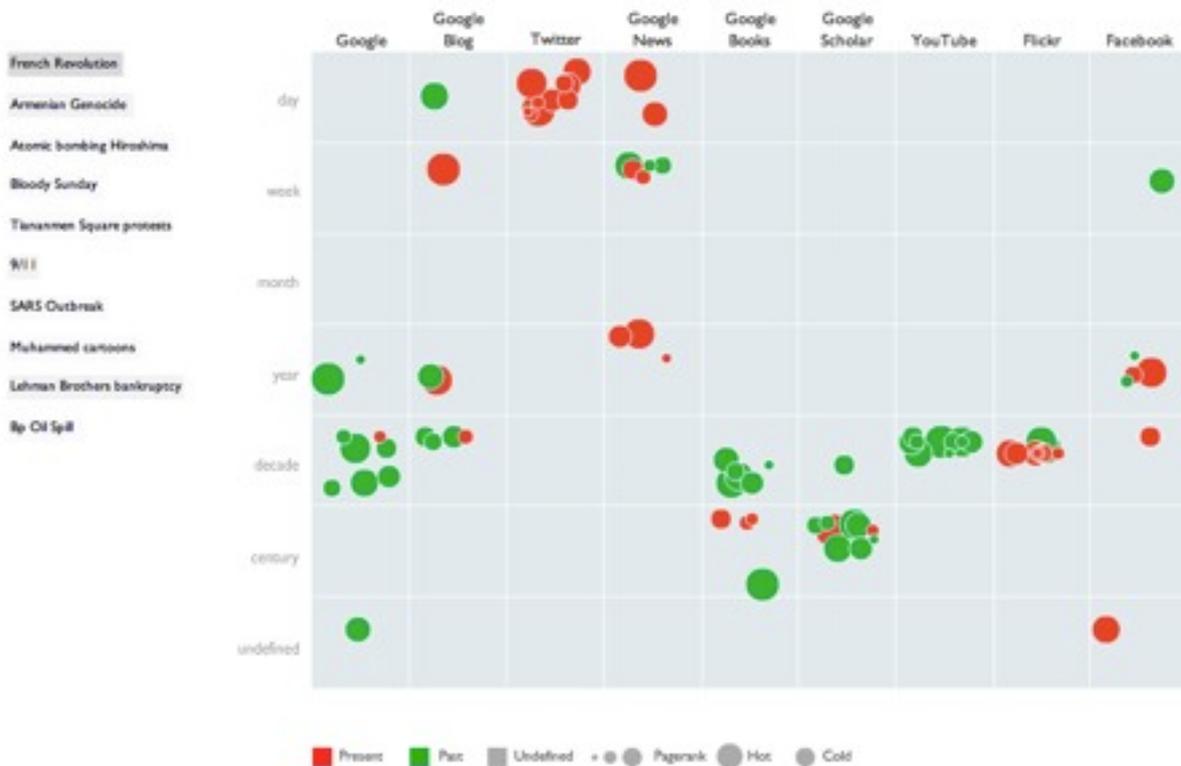


Fig. 2 Reworking of the French Revolution across Web platforms

Historical controversies Now

Querying historical controversies in dominant devices and platforms, the question we ask is what kind of history are we accessing on each device?

[Read more](#)

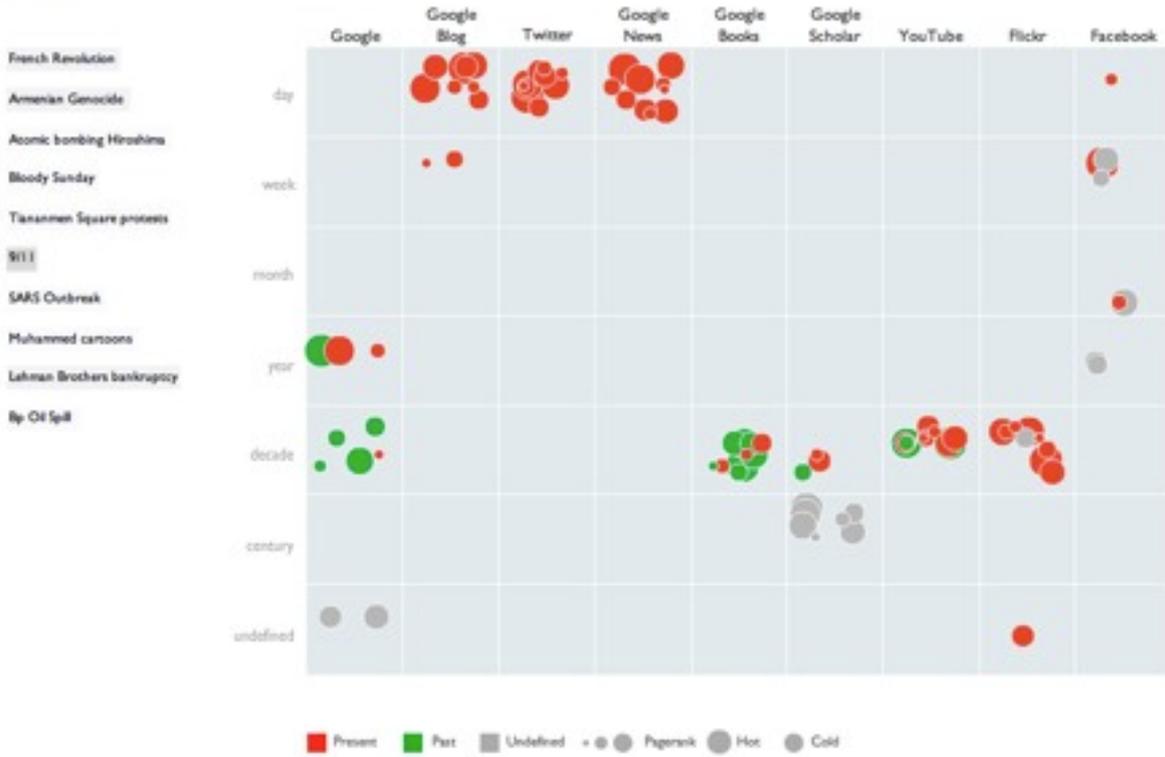


Fig. 3 Reworking of the 9/11 across Web platforms

The Lehman Brothers bankruptcy sources (Fig. 4) are in stark contrast with 9/11 ones: the overall time dynamic of the bankruptcy is in the past with very few daily or weekly reworkings. However, this is the historical controversy that is disputed on Google Scholar and entirely past and settled on Facebook. Flickr sources (although in the past) contest the events more than YouTube which is a pattern found across the histories.

“Historical Controversies Now” show that different platforms have specific dynamic of remembering. Twitter and Google News actively rework and make present various historical events. The research team observed that the present actualization happens in various ways, for example, through personalization of the historical event (reading a history book, visiting a historical site, listening to a song), metaphorically invoking the event, or through staging a historiographic debate. Google Books and Scholar are on the other side of the spectrum with past sources and low contestation of the historical controversies. Surprisingly, Google Search falls also in this category of retrieving sources and documents (Wikipedia articles, history sites) without reworking them. The research team concluded that for the social Web platforms (YouTube, Flickr, Facebook) the actualization of the past depends considerably on the event: Facebook reworks the Armenian

Historical controversies Now

Querying historical controversies in dominant devices and platforms, the question we ask is what kind of history are we accessing on each device?
[Read more](#)

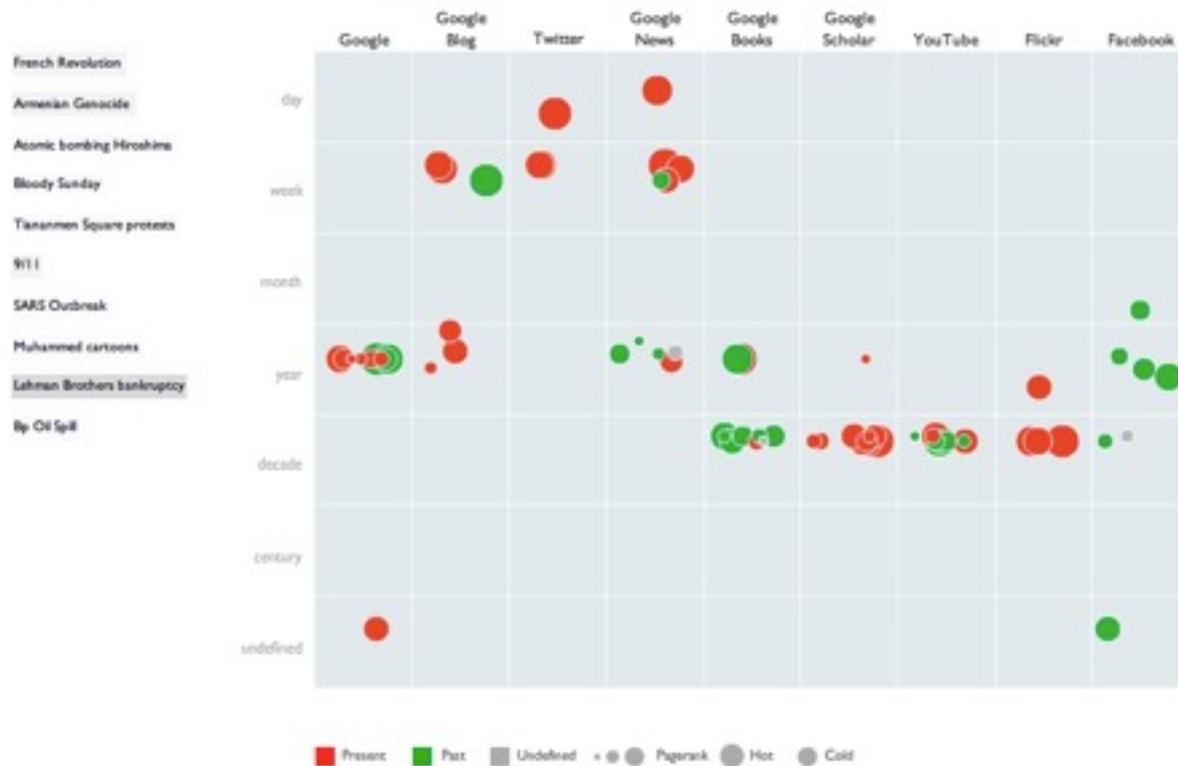


Fig. 4 Reworking of the Lehman Brothers bankruptcy across Web platforms

Genocide and settles the Lehman Brothers bankruptcy, for example. However, Flickr sources were found to contest the events more than YouTube which is a pattern found across the histories.

Comparing between the historical events, the DMI researchers found that some histories are more actively reworked and negotiated as controversial. This is very explicit in the ways the Web platforms handle the French Revolution and the Armenian Genocide. Most of the sources for the French Revolution (apart from the ones on Twitter and Google News) give an account or document the historical event itself, without explicitly reworking it. On most of the platforms, the French Revolution it is not actualized as controversial. It is reworked metaphorically or it is personalized (somebody reading a book on the French Revolution). In opposition, the Armenian Genocide is very actively mobilized as a controversial event, largely debated and contested. Therefore, the project not only maps the dynamic of remembering for several past events on different platforms but also engages with the mechanisms through which the histories are reworked and made controversial.

4.2 “Neutral or National Point of View? Wikipedia’s language versions as cultural reference”

After engaging with a case study that examines the ways different Web platforms (dominant for a particular sphere) negotiate remembering and how they actualize and contest past events, the article will zoom into a specific platform for an in-depth understanding of the mechanism by which they rework and retell memory. The next case study that would be presented and discussed here is a project undertaken by Emina Sendjarevic, a MA student of “New Media & Digital Culture” program, University of Amsterdam. “Neutral or National Point of View? Wikipedia’s language versions as cultural reference”¹¹ tracks a Wikipedia article through its different language versions to find and explore significant ‘deviations’ from the English Wikipedia entry (seen as the example of incorporating the ‘neutral point of view’ - the standard for ensuring qualitative encyclopedic content in Wikipedia). The Wikipedia entry is on ‘Srebrenica’ which refers to the killings of more than 8000 Bosnian men and boys by the Army of the Republic of Serbia in the United Nations protected zone around Srebrenica in 1995. According to the International Criminal Tribunal for Former Yugoslavia and the International Court of Justice the killings in Srebrenica constitute a genocide (ICTY 2004; ICJ 2007).

The project, therefore, traces a highly contested, controversial, and emotionally charged past event and its rework in six language Wikipedias: Serbo-Croatian (the official language of former Yugoslavia), Serbian, Bosnian, and Croatian (the national languages that were introduced in the corresponding countries after the Yugoslavian wars of 1991-1995), English (the international perspective) and the Dutch language Wikipedia entry since the UN Protection Force stationed in this area was from The Netherlands (the ‘Dutchbat’). Methodologically, Sendjarevic incorporates content analysis on the “Table of Contents” (title, subtitles, headers) of the language Wikipedia entries on ‘Srebrenica’ to detect distinctive framing of the event and its subjects; outlink analysis to determine the verifiability of the content (the Wikicommons outlinks, geo-hacks and the general Wikipedia pages were excluded from the outlinks count); and contributors evaluation. Contributors were listed using the DMI Wikipedia Edit Scraper and IP Localizer tool and the Raw Text to Tag Cloud Engine and The TagCloud Generator tools were then used to scale and visualize the contributors into a tag cloud. The settings of the tools were specified to exclude ‘strings smaller

¹¹ “Neutral or National Point of View? Wikipedia’s language versions as cultural reference” is being reworked into an article and is expected to be published later this year.

than 0 characters' and 'words with frequency smaller than 3', therefore users that contributed to the entries less than three times were not included in the sample.

The findings from the content analysis signal a clear differentiation of the framing across the language Wikipedias. The Bosnian and Croatian entries frame the killings in Srebrenica as 'genocide' while the Serbian and the Serbio-Croatian version refers to it as a 'massacre'. Sendjarevic notes explicitly that the articles in Serbian and Croatian were created by one contributor who deliberately chose to frame it 'massacre' in the Serbian entry and 'genocide' in the Croatian one. The English Wikipedia defines the events in Srebrenica also as a 'massacre' (following the Wikipedia guidelines for controversial subjects that should be named on the basis of 'recognizability'¹²) while the Dutch article use a military term 'the fall of Srebrenica'. Distinct reworks occur in the number of the victims: the English Wikipedia entry is the most exact; the Bosnian and Croatian versions provide an estimate while the Serbian and the Dutch article play down the numbers. The framing of the responsibility and the subjects involved is reworked as the following: the English, Bosnian, and Croatian language versions highlight the Serbian responsibility; the Serbian article uses the name of the military operations ('Operation Krivaja 95'; 'Operation Stupcanica 95') while the Serbio-Croatian and the Dutch version neutralize the event as the 'fall of Srebrenica'. Framing the controversies surrounding Srebrenica happens also differently across the language Wikipedias: the English version recognizes several contestations around the event ('Possible widespread racism among the Dutch peacekeepers'; 'Greek voluntaries controversy'; 'Dispute regarding the Serb casualties around Srebrenica', etc.); the Bosnian and Croatian article highlight the 'denial of genocide' while the Serbian and the Serbio-Croatian versions project 'Criticism of the official versions', 'Alternative visions of the events, revisionism, and conspiracy theories'.

Concerning the number of the outlinks, Sendjarevic found the Serbian Wikipedia entry has the most outlinks to outer sources among the ex-Yugoslavian language articles. However, it is labeled 'biased' while the other entries have less outlinks and their validity is not disputed. Contributors community was also found to be specific for every language Wikipedia: the power editors (Styilia et al. 2007) do not contribute to other ex-Yugoslavian articles besides their own. They are sufficient within heir own language specific article and tend not to edit even the Serbio-Croatian entry that symbolizes their common Yugoslavian past. However, a small group of Serbian and Bosnian editors tend to contribute to the English version (Appendix A). The anonymous

editors, on the other hand, edit the entries across languages (except for the Dutch entry). The rework and the contestation of the event occurs through the anonymous edits (Appendix B) while the Dutch language article is compiled at most by IP-addresses located within The Netherlands. Sendjarevic concludes that given the distinctive reworking of Srebrenica killings through the language Wikipedias, different national points of view have been explicated. She highlights various strategies that have been incorporated to do so without surpassing the Wikipedia guidelines: creating a language version of an article (Bosnian, Croatian, Serbian) for essentially the same language (Serbo-Croatian); following the Wikipedia rules of framing controversies in relation to 'recognizability' (the choice of the English version power editor to define 'Srebrenica' as 'massacre' not 'genocide'); using many sources to validate the rework and forking the article when its reliability is disputed (the case of the Serbian entry). None of these strategies is applied when the article represents a community that shares the same national perspective. In the case of the Dutch entry of 'Srebrenica' contestation of the event has not been found.

This case study has been incorporated into the argument to explicate the mechanisms by which memory could be mobilized and reshaped using the specificity of a Web platform. The language Wikipedias enable different past narratives that could be brought to serve a broader national / cultural agenda. The means by which this is done lay in what is enabled by Wikipedia: specific language entries, citations for verification, recognizable framing in the cases of controversy, anonymous edits and forking.

Possible limitations of the projects could be related to the use of Web's data. The case studies follows Parikka & Sampson's call for moving away from the solely representational analysis in media and cultural studies focused (mostly) on the effects of the medium. However, the empirical collection based on Web's data is subjected to change over time (online platforms also continuously update their norms and rules). The data on the Web is in a constant flow thus we should rethink what constitutes an online research. The expectation from the scientific work that you can always go 'back' and obtain similar results is not applicable for the Web due to the (ever) changing 'nature' of its data. The specificity of the software tools (what they enable and constrain) used for capturing, analyzing, and visualizing the data also plays an important role in the analytical process. Therefore, the digital methods (used in medium - specific projects presented in this article) do not output set of results (objective/right or wrong) but 'patterns' (based on results that change overtime).

5. Conclusion

The article has introduced a medium-specific approach in studying the dynamics of remembering and forgetting on the Web's (social) platforms. It started with a brief overview of the dominant modes of research concerning this dynamic - the limitations of the online archives to utterly preserve the content of the Web (perceived as a source of knowledge); the abundance of traces and memory of the users (and the implications for privacy); and the impact of the Web on what and the way we remember. The medium-specific analysis enables studying the content with a focus on the ways the Web's platforms rework, reshape and modulate the memory of past events. By presenting and discussing two case studies ("Historical Controversies Now" and "Neutral or National Point of View? Wikipedia's language versions as cultural reference") we were able to map the dynamics of remembering and forgetting across platforms (and Web spheres) and within a platform.

The analysis concluded that different platforms have specific dynamic of remembering. Twitter and Google News actively rework and make present various historical events. Surprisingly, Google Search falls in this category of retrieving sources and documents (Wikipedia articles, history sites) without reworking them. The social Web platforms (YouTube, Flickr, Facebook) actualization of the past depends considerably on the event: Facebook reworks the Armenian Genocide and settles the Lehman Brothers bankruptcy. However, Flickr sources were found to contest the events more than YouTube which is a pattern found across the histories.

Another finding is attributed to the mechanisms by which the reworkings are undergone. Especially in the case of Wikipedia, they are exclusively enabled by the platform itself: specific language entries, citations for verification, recognizable framing in the cases of controversy, anonymous edits and forking. This finding could be related to Barbara Misztal's consideration of constructing memory 'bottom up'. In *Theories of Social Remembering* Barbara Misztal refers to Halbwachs concept of collective memory through the possibility of its controlled content (Misztal 2003: 53). There are theories that refer to the way the past is selected, constructed and applied 'from above' to exercise power and authority. However, Misztal introduces also the theory of possibility to construct the memory from 'bottom up' as starting with the local and particular and building gradually up (Misztal 2003: 61). Social groups can mediate their own version of the past as a 'counter' or 'unofficial memory' that can shape their identity. Wikipedia analysis has shown that the mechanisms to do so are embedded within the medium.

7. References

- G. Blome and G. Wijers (2010). "Visibility, Distribution and Memory through Networks and Collaboration" in *Archive 2020, Sustainable Archiving of Born-Digital Cultural Content*. Virtueel Platform: Amsterdam.
- N. Brügger (2009). "Website history and the Website as an object of study." *New Media & Society* 11.1-2: 115, pp. 290.
- W. H. K. Chun (2010). *Programmed Visions: Software and Memory*. Cambridge, MA: MIT Press.
- (2008). "The Enduring Ephemeral, or the Future is a Memory." *Critical Inquiry* 35.1: 148-171.
- Critical Art Ensemble (1995). "Utopian Promises – Net Realities." Hamburg: conference proceeding of *Interface 3*, pp. 146.
- S. Leong, T. Mitew, M. Celletti, E. Pearson (2009). "The Question Concerning (Internet) Time." *New Media & Society* 11.8: 1267–1285.
- A. Maj and D. Riha eds. (2009). *Digital Memories. Exploring Critical Issues*. Oxford, UK: Interdisciplinary Press.
- B. A. Misztal (2003). *Theories of Social Remembering*. Maidenhead: Open University Press.
- J. Parikka and T. Sampson (eds.) (2009), *The spam book: On viruses, porn, and other anomalies from the dark side of digital culture*, Cresskill, NJ: Hampton Press
- A. Reading (2008). "The Globytal: Toward an Understanding of Globalised Memories in the Digital Age." *Digital Memories. Exploring Critical Issues*, eds. Anna Maj and Daniel Riha. Oxford, UK: Interdisciplinary Press: 81-95.
- R. Rogers (2009). *The End of the Virtual: Digital Methods*. Amsterdam: Amsterdam University Press.
- "Digital Methods Initiative." <http://www.digitalmethods.net/>, retrieved 5 October 2010.
- S.Schneider and K. Foot (2004). "The web as an object of study." *New Media and Society* 6.1 (2004): 114-122, pp.119
- V. Mayer- Schonberger (2009). *Delete: The Virtue of Forgetting in the Digital Age*, Princeton University Press.
- E. Sendjarevic (2010). "Neutral or National Point of View? Wikipedia's language versions as cultural reference." DMI final paper.
- M. Stevenson (2009). "The Archived Blogosphere, 1999-2001." in *Changing Cultures: Cultures of Change*, draft conference paper.
- B. Stvilia, M.B. Twidale, L.C. Smith, and L. Gasser (2008). "Information quality work organization in Wikipedia." *Journal of the American Society for Information Science and Technology*, volume 59, issue 6, pp. 983-1001, at <http://www3.interscience.wiley.com/journal/117919271/abstract>, accessed 20 December 2010.

D. Sullivan (2009), “Google’s Personalized Results: The ‘New Normal’ that Deserves Extraordinary Attention” (published online <http://searchengineland.com/googles-personalized-results-the-new-normal-31290>, 2009), retrieved 8 February 2010.

G. Taddeo (2009). “The Blended Boundary between the Individual and Social Memory in the 2.0 Web Era.” *Observatorio (OBC*) Journal* 11: 115-131.

R. Walker (2010). “Cyberspace When You’re Dead.” *New York Times Magazine*: 2010, http://www.nytimes.com/2011/01/09/magazine/09Immortality-t.html?_r=4&pagewanted=all, retrieved 12 January 2010.

Esther Weltevrede (2011). “Atemporal Engine, and Why a Query’s Date is Important.” DMI Winter Conference paper.

ICJ, 2007. “Application o the Convention on the Prevention and Punishment of the Crime of Genocide (Bosnia and herzegovina v. Serbia and Montenegro,” *ICJ*, (26 February), at <http://www.icj-cij.org/court/index.php?pr=1898&pt=3&p1=1&p2=3&p3=1>, accessed 17 December 2010.

ICTY, 2004. “Address by ICTY president Theodor Meron, at Potocari Memorial Cementery,” *ICTY*, (23 June), at <http://www.icty.org/sid/8409>, accessed 17 December 2010.

Wikipedia, 2010k. “Article titles,” at http://en.wikipedia.org/wiki/Wikipedia:Article_titles, accessed 22 December 2010.

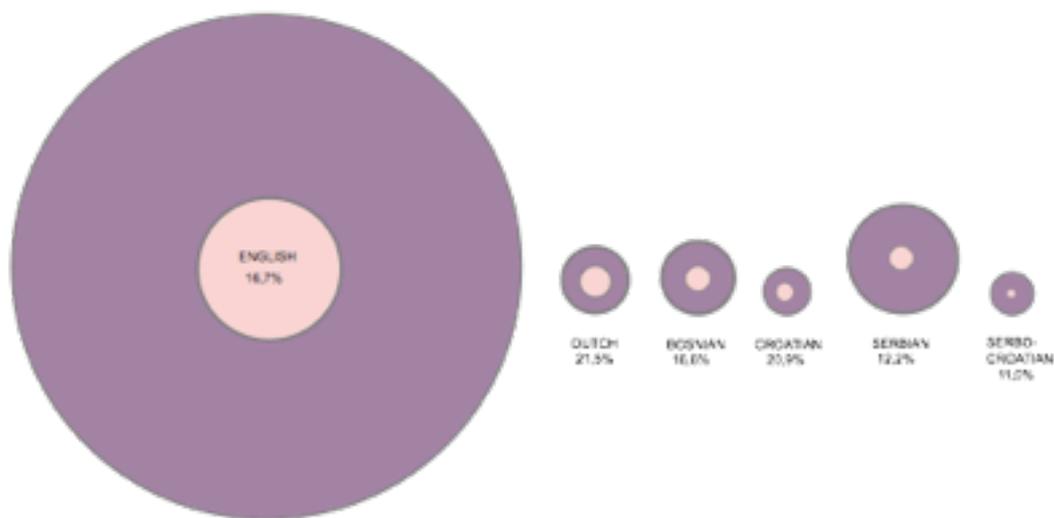
<https://wiki.digitalmethods.net/Dmi/OldControversiesNow>, retrieved 6 May 2010.

Frequency of edits by (anonymous) users across language versions

Which articles about the Srebrenica genocide are edited the most? How much is edited anonymously?

Method: Data taken from the Wikipedia Edit Scraper and IP-localizer, visualisation by Bubblelines.

Conclusion: The English version has the most edits. Remarkably the Serbian contains the most edits, although the content is the least extensive. Also it seems that the Dutch and Croatian have the most anonymous edits. Overall the portion of anonymous edits is low (around 13%).



Frequency of usernames by edits across language versions

Did a large group of users edit the content or only a few? Who are the users that contribute/edit?

Are articles written in ex-Yugoslavian languages edited by the same users?

Method: Data from Wiki/Pedia Edit Scraper and IP-localizer, visualised by Tag Cloud Generator

Conclusion: In the English article a group of heavy editors/contributors can be found. Here also users that edit the Bosnian and Serbian articles are found. Remarkably the Dutch, Serbo-Croatian and Croatian articles are primarily edited by one or two users. The Serbo-Croatian article is for a big part co-authored by Bots. The Dutch article has its own set of users that edit.

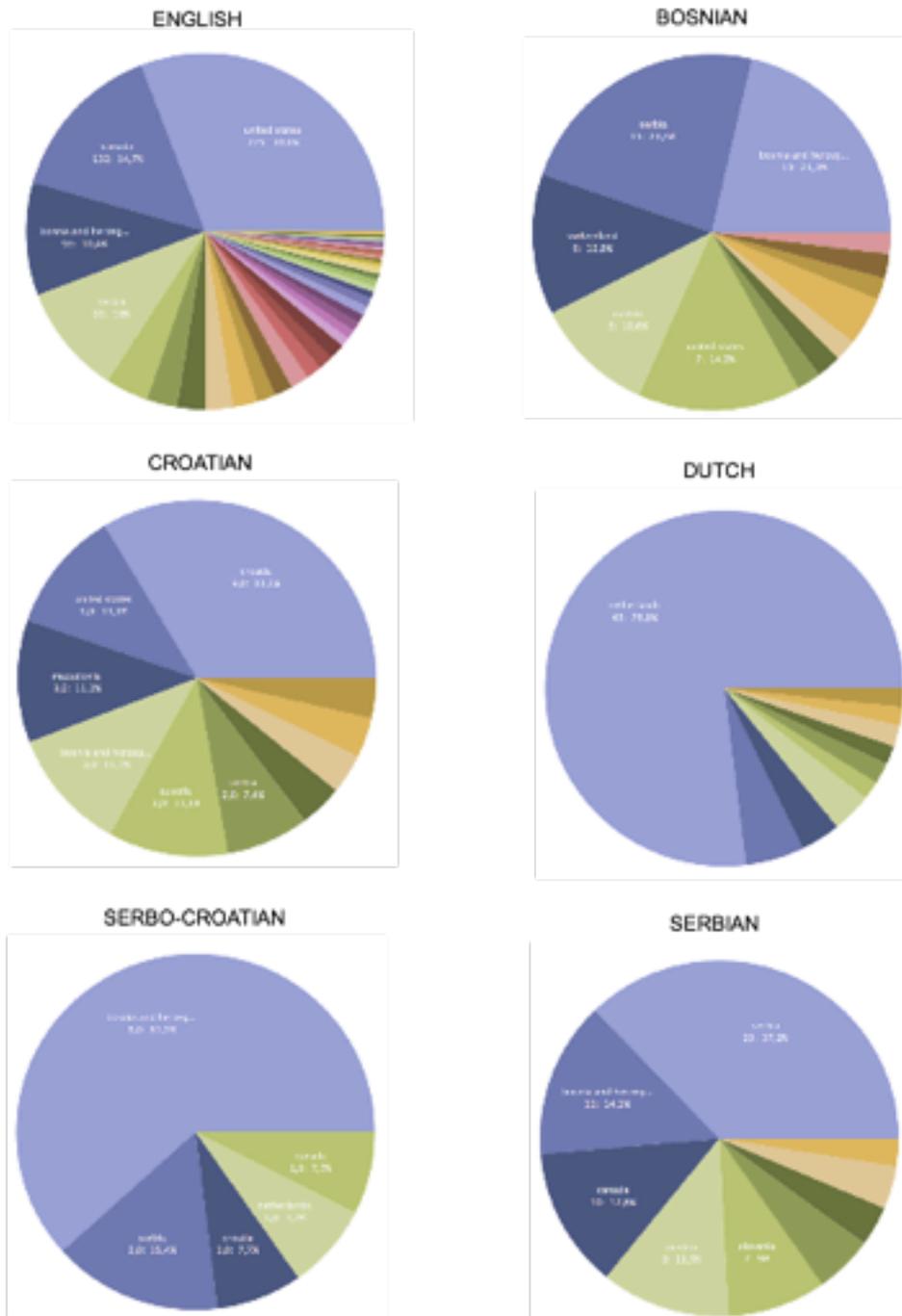


Where do anonymous edits come from?

Frequency of countries anonymous edits have been made from for the Wikipedia article about the Srebrenica Genocide

Method: Data from the Wikipedia Edit Scraper and Localizer tool, Visualisations from ManyEyes.

Conclusion: The ex-Yugoslavian countries (Bosnia, Serbia and Croatia) are editing each others Wikipedia articles about the Srebrenica genocide, or at least anonymously. Anonymous edits in the English article are the most internationally divers. The Dutch article is mostly edited from Dutch IP-addresses.



Appendix B

Visualizations: Emina Sendijarevic