

Against Search—Towards A New Computational Logic of Media Accessibility

Pelle Snickars

Let's start off with one of the most compelling questions of our time: what does it mean to be human in the digital age? Well, one overwhelming challenge facing us all is having digital access to more information, data and knowledge than any previous generation of humankind. A burden perhaps—at least for some. But for the majority of us, a blessing. The often invoked libertarian *information-wants-to-be-free* paradigm not only insists on free flow of data. All these bits and bytes in the digital domain has to be organized and found, which needless to say is the underlying rationale for the most successful web behemoth of all. Suffice to say, we all live with an increased screen attention (of various sizes), and giving computers (and their mobile clones) textual and haptical commands has also become a ubiquitous normality. Access to whatever we want literally lies at our fingertips; information is there somewhere waiting—and the question are always where to look. So, you search.

Ever since Google introduced its white and clean search box interface in the late 1990s—Internet Archive crawled the site for the first time in mid November 1998—the blank frame has been waiting for input.¹ During the last decade this peculiar type of white box has become the new search *default*, especially within the information retrieval sector *par excellence* at archives, libraries and museums. “Search the Collections”, is the standard phrase awaiting every online user, implying a more or less vague notion that one already needs to know what one came for. Users are, of course, experienced since surfing the web basically means searching it. Subsequently, the notion of ‘search’ is key for the digital domain in general, and the web in particular. Understanding Google, Steven Levy notes, is trying to “grasp our technological destiny.”² From a more scholarly perspective, ‘Search Studies’ is on the brink of developing into an academic field; ‘search’ is, after all, *the* primary human-computer interaction mode. Mining search patterns and optimizing the engine is what Google and other search companies does on a daily basis, and through online ‘search’ events IRL, like the spread of flues, can increasingly be anticipated. Search *per se* has in many ways somewhat paradoxically become the answer to questions asked.

To diagnose the cultural logic of online search is, naturally, a vast topic—ranging from the omnipresent potential of Google analytics to the critique of the ‘googlization of everything’ and unfiltered initiatives like Scroogle. Being coded and technical by nature ‘search’ remains highly complicated, with constant upgraded algorithms exploiting the link structure of the web. Since studying tech infrastructures is a blind spot for media studies, complexities are particularly striking from this perspective. Accessibility to various media content in an age characterized by dynamics and volatility is, however, regulated by notions of search, and therefore it remains essential to analyze and grasp how and why ‘search’ has become so important.

During the last decade the notion of search has also been challenged by new and alternative computational modes of accessibility, which is yet another argument why ‘search’ needs to be taken seriously (and, admittedly, few would argue otherwise). Tags, folksonomies, or social tagging are, for example, new transformative web based practices and methods to annotate and categorize information and media content in an effort to collectively classify, tease out and find data in other ways than simply through the mantra, ‘search the collections’. Online *browsing* is, of course, a widely used option, as well as simply ‘clicking’. On YouTube—the quintessential new digital ‘archive’—one textual search is often enough, and then tags and linked videos leads the user into a streaming vortex of differentiated media. Context of content is often fleeting and arbitrary; odd juxtapositions norm rather than exception, and material regularly detached from its place of origin. Clicking rather than searching, thus, becomes an epistemic way of locating and perceiving media material, often in unintended ways. Usage resembles that of walking around in (weird) open library stacks, even if the much appraised digital ‘openness’ on the net in general, and on web 2.0 platforms in particular, always remains modulated on a protocological basis. A web browser is, after all, a translator of code and an interpreter of digital data that profoundly shapes user experiences. Then again, from a strict computer-science perspective, user generated and participatory platforms like YouTube are nothing but databases. Still, in any given cultural context, surfing onto a platform and watching a video at, say, YouTube obviously entails more than that. From a media studies perspective it is therefore debatable whether we ‘watch databases’ only (Lovink), or claims that there is ‘no content, only data and other data’ (Galloway & Thacker), has much relevance in regards to YouTube, or for that matter other cultural heritage or social media sites.³

Nevertheless, given the sheer size of contemporary online media collections—from the vast information repositories of data at Wikileaks or The Pirate Bay, to billions of UGC on YouTube and Flickr, or for that matter the 20 million digitized heritage objects at the Library of Congress—simply having a look what’s inside the digital ‘archive’ is no longer possible. However, the contemporary ‘flood of information’ is, by no means, new. On the contrary, libraries and archives have during the last century repeatedly complained over way too many books and documents. The major difference, today, is that in digitized form such material can be analyzed collectively as major cultural sets rather than on a singular basis only. Singularity works for analyzing the particular. But the general is arguably more interesting, and often of greater importance. Hence, massively linked data has nowadays the potential to reveal new human patterns that hitherto remained invisible. The notion of a particular ‘search’, then, is not the answer to the more or less *infinite* digital archive.

New Modalities of Access

This article—borrowing its title from a recent blog post of Lev Manovich⁴—tries to map out, explain and understand, as well as situate and critically examine new search modalities within a larger framework of information retrieval in general, and alternative forms of media archival accessibility in particular. In short, I will argue that new forms of different computational logics should increasingly be deployed in order to facilitate access to deep data as well quantitative surface data in both *web n.0* media collections and at more traditional digital archives and libraries currently being coded online. Archives and libraries are, however, conservative by nature; not even code changes that. Leaving aside the fact that memory institutions regularly use web 2.0 services, at an institutional level archival material still needs to be *found* by way of a distinct, or even advanced search (whatever that means). So, rather than taking advantage of new digital media—and in a sense ‘follow the medium’—the notion of ‘search’ is still embedded in traditional forms of archival access.

Open data, freely distributed APIs and the increased sharing of networked data has, as is well known, led to new ways of distributing information (and knowledge). Then again, even if online access is surrounded by buzzwords as ‘democracy’, ‘free culture’ or ‘networked society’, accessibility is still regarded as a rudimentary and, quite simplified question of search, not the least within the cultural heritage sector. But, there

are of course new ways of examining digital content through code and technological input rather than sheer human agency. New modes of content based image retrieval systems or information visualizations are for instance illustrative cases. An example of the former, related to Google's Similar Images, can be found at Europeana's ThoughtLab on heritage data, which presents an image search demo that scans 70,000 images from Europeana data providers, and allows one to "perform image search based on a content-based retrieval technology using ... visual descriptors to decide on the similarity of images."⁵ IBM's "Many Eyes" and its data visualization tools, is an example of the latter, where users can upload data and then produce graphic representations for others to view and comment. Data visualization is, often, understood as the mapping of digital data onto a visual image, and so called 'info vis' of large-scale collections of non-numerical information—as files or lines of code in software systems or bibliographic databases—has been a major contemporary trend during the last years. Related are also ways of scrutinizing "big social data"—that is, the ability to analyze data in scale within social networks, thus, "enabling a range of intriguing and useful applications that can plug into social media networks and make use of the knowledge inside them", as *ZDNet* put it in May 2011.⁶

However, as Richard Rogers keeps reminding, there remains an ontological distinction between the 'natively digital' and the 'digitized', that is, between digital objects and content 'born' in new media online, as opposed to, for example, scanned cultural heritage that have migrated into the digital domain. Based on code, the former material can be analyzed in myriad of ways, whereas the latter often takes the form of a representational image file (to be searched for). In short, *digitized* material is not 'digital'. A new emerging computational logic of media accessibility, then—whether centered on haptic touch, information visualization, graphic similarities or algorithmic search—needs to take this particular polarization of the 'natively digital' and the 'digitized' into account. Still, it goes without saying that various forms of "digital methods" ought to be used when approaching major social media collections or a heritage in bits online.⁷

From a 'search' and media dynamics perspective, what's apparent today are on the one hand new emerging search patterns, especially in relation to the cultural heritage sector as well as models of algorithmic search, and on the other how archives, libraries and museums are currently working with new ways of (re)presenting their digitized collections, where rethinking design and archival interfaces is one way of altering and

finding new modes (and nodes) of accessibility. Recently, the British Library, for example, released a “19th Century Historical Collection App” with more than a 1,000 rare books with titles in fields like travel writing, natural history and philosophy. Users of the app have, of course, the ability to search the included books, but the individual items are predominantly viewed as high-resolution scans. Hence, one might argue that the iPad’s touch screen *per se* functions as a kind of browse default to literally *handle* the included material. Perceiving these old books in digitized form, then, not only becomes a matter of ‘search’ or reading, but more so in terms of haptic treatment, fondle and physical engagement—all smoothly executed with a gentle touch. According to a library press release, the app is said to take “advantage of the form and function of iPad, bringing a renewed sense of wonder to the discovery and enjoyment of antiquarian and historical books.” As such, the app represents the latest landmark in the British Library’s progress towards its “long-term vision of making more of its historic collections available to many more users through innovative technology.”⁸ Another similar example of a sort of haptic archival immersion is the The New York Public Library’s *Biblion* app, centered around the official corporate records of the 1939/40 New York World’s Fair. The app contains some 700 documents, images, films and audio material, and has been described as “one of the slickest media consumption experiences” that have yet been released for any tablet.⁹ According to NYPL, the app takes the user “literally into the Library’s legendary stacks, opening up hidden parts of the collections and the myriad storylines they hold and preserve.”¹⁰

Aspects of haptic immersion, then, is one new fascinating computational logic of media accessibility within the memory sector slowly to emerge and mainly driven by new forms of tablet hardware and ‘the app revolution’. Still, these apps also emphasize how technology always regulates access to the past, especially given that they are curated instances of selected entries into collections. It is essential to keep in mind when trying to map out new search modalities, that underlying code regulates what can—and can’t be done. Even if it goes without saying, systems chosen for search and accessibility will always determine the output of presented information, as well as research performed. And again, needless to say, these are important epistemological issues. Searching the web is, as we all know, critical to the ability of using the net, and whoever controls search engines has enormous influence. At present Google shapes and regulate what “we read, who we listen to, and who gets heard. Whoever controls the search engines, perhaps, controls the internet itself”, as James Grimmelman has put

it.¹¹ Technology is, and has never been neutral. The suspicion that our writing tools are always working on ‘our thoughts’, was after all raised by Nietzsche already in the 1880s. More than a century later, after Biblical efforts of arranging and cataloguing, describing and digitizing cultural heritage content the pattern remains—or as Lev Manovich recently stated: “hierarchical classification systems used in library catalogs made it difficult to browse a collection or navigate it in orders not supported by catalogs.” Walking from shelf to shelf one had to follow a classification system based on subjects, “with books organized by author names inside each category.” Taken together, these distribution and classification “systems encouraged 20th century media researchers to decide before hand what media items to see, hear, or read.”¹² Today, however, the situation has changed and it is no longer impossible to imagine navigating—in one sense or another—through *all* collected material of a given topic, which, arguably, is an insight *in itself* that makes a difference.

More Data is Better Data

Under the heading “Humanities 2.0” *The New York Times* ran a series of articles during the winter of 2010/11 on how digital tools are changing human scholarship. According to one of the pieces, members of new generation of “digitally savvy humanists” didn’t look for inspiration anymore in the next “political or philosophical ‘ism’” but rather wanted to explore how digital technology as an accelerating force was changing the overall understanding of the liberal arts. New methodologies, powerful technologies and vast stores of digitized materials “that previous humanities scholars did not have” acted as a revisionist call of what human research is all about.¹³

Given the conservative culture of scholarship in general, and humanistic research in particular, the basic arguments in these articles were striking, not the least since they articulated the increasing role that computerized technology plays for humanistic research (whether it wants it or not). If the computer is the cultural machine of our age, to invoke the notion of Peter Lunenfeld, then the same goes for research. The field of digital humanities is rapidly picking up speed—often closely linked to the cultural heritage sector—and the discursive idea of the lone scholar, working in isolation with his or her own archiving solutions, will all likely (at least in due time) fade away. As the report, *Our Cultural Commonwealth* stated already in 2006, humanistic researchers and users of “massive aggregations of text, image, video, sound, and metadata will want

tools that support and enable discovery, visualization, and analysis of patterns; tools that facilitate collaboration; an infrastructure for authorship that supports remixing, recontextualization, and commentary—in sum, *tools that turn access into insight and interpretation.*¹⁴

To be honest, we are not quite there yet. Still, there are many examples of a new media dynamics that involves upgraded modes of archival accessibility. One successful archival project, for example, is the “Transcribe Bentham—A Participatory Initiative”, under the auspices of the Bentham Project at University College London, which aims to produce new editions of the scholarship of Jeremy Bentham.¹⁵ Transcribe Bentham is, in short, an open source and participatory online environment launched to aid users in transcribing 10,000 folios of Bentham’s handwritten documents, and as such it has been invoked and discussed at length within the digital humanities. Speaking at a conference during the summer of 2010, Melissa Terras, for example, stated that crowd sourcing and the harnessing of online activity “to aid in large scale projects that require human cognition—is becoming of interest to those in the library, museum and cultural heritage industry, as institutions seek ways to publically engage their online communities, as well as aid in creating useful and usable digital resources.”¹⁶

If haptic entries and touch screen interfaces on tablets is a minor current archival trend, then, a really major trend is heritage accessibility on a massive scale via large cultural data sets, as the example of Bentham’s ten thousand handwritten documents show. Google has naturally paved the way for this overall change of perspective and scope, and have now digitized more than 15 million books—many of which belong to the public domain. According to a recent post on the Inside Google Books blog, 150,000 stem “from the 16th and 17th centuries, and another 450,000 from the 18th century.”¹⁷ Leaving aside the fact that Google Book Search has ran into copyright problems, their digitization efforts have shown that massive scan projects, on an hitherto unthinkable scale, can very well be undertaken. In addition, as Robert Darnton has recently pointed out, “it is too early to do a postmortem on Google’s attempt to digitize and sell millions of books, despite [current legal problems.]. Google Book Search may rise from the ashes, reincarnated in some new settlement with the authors and publishers who had taken Google to court for alleged infringement of their copyrights.”¹⁸ Data from the scanned books are, in fact, already being widely distributed. With Google’s Ngram Viewer, for example, it’s possible to visualize the rise and fall of particular keywords across these scanned millions of hundred-year-old

books. The respective term ‘archive’ and ‘database’, for example, generates almost zero interest before the late 1960s—then suddenly books start filling up with these very notions.

Mining textual archives and visualizing the results in various ways, is, alas, one contemporary strategy of moving beyond the white search box. ‘Text’, however, largely remains the organizational mode of accessibility, and besides still make up the bulk of content in ‘new’ digital archives. Then again, the long awaited celestial jukebox of all known media can’t be discarded as a fancy thought anymore. On the contrary, it’s already a reality through file sharing or legal P2P sites as, for example, Spotify—which presently boasts access to approximately 15 million songs, and where social recommendations through ‘friends’ increasingly points to alternatives access modes for media. Everything that can be digitized—will be digitized, the catch phrase once went during the 1990s, and a pertinent question is what such a claim actually implies. What practical consequences will it lead to on an institutional level, for instance? Apparent is that a new cultural logic of media accessibility has began emerging due to the sheer size of digital collections, quite different from traditional and analogue search modalities.

Of course, browsing a library catalogue has its particular media history. Ever since institutional heritage catalogues were transferred to digital formats during the 1970s and 1980s similar search concepts and notions as today have been employed. If these catalogues were once browsed manually by hand, often with the researcher leaning over giant stacks of index cards, in computerized form typing a command gave way for a new logic of access. Through input of textual commands on a computer screen, (re)searchers started to locate metadata—that is, information on and about the searched material, whether it being a book, an image or a film—through highly subjective textual input. Today, with a cultural heritage increasingly being digitized, transformed into and represented as data, a similar mode of logic structures access, yet at the same time the potential is also there to navigate in completely new ways.

If user generated content online has experimented with new classification systems during the last decade, institutionalized digitized heritage is still, however, basically accessible and found through ways of ‘search’ in online galleries. “See 30,000 items from our collection”, as the British Library puts it online. “Ideas and inspiration can be found within the more than 15 million items on Europeana”, according to the major heritage portal europeana.eu. Hence, searching a database of cultural items literally means executing a discrete set of commands: decide what to search for, browse the

obtained metadata, and—if the object in question has been digitized—get access to an imagistic representation, which can hopefully be downloaded for research or re-use. Europeana, for instance, boasts that it enables people to “explore the digital resources of Europe’s museums, libraries, archives and audio-visual collections. It promotes discovery and networking opportunities in a multilingual space where users can engage, share in and be inspired by the rich diversity of Europe’s cultural and scientific heritage.”¹⁹

Then again, describing the above is merely stating the obvious. The bigger picture is how you actually search (or browse) 15 million cultural items—whether in a classical or *web n.0* fashion? Where do you start, and according to what principle are you trying to make sense of *all* material potentially at your disposal? From a researchers perspective, trying to relate to some scientific method, the task is simply put, totally impossible. You just can’t cope with so many items. No human user can do that—only a computer (or a network of them). There is no way of making analytical sense of 15 million cultural items. New digital dynamics of interlinkage, participatory tagging or algorithmic search rather points towards the need for new mode(l)s of approach to such vast collections.

As a consequence, digital humanists and researchers working with major cultural data sets have begun to pose questions as to whether new digital archives, understood in a broad sense as massive collections of data, can be analyzed and searched *at all* in traditional ways. If humanist scholars previously worked by personally extracting data from archives, gleaning bits and pieces often found haphazardly, the millions of items in, for example, Europeana seems to call for, or at least imply a new practice (as well as theory) of humanist research, involving the very machines that transformed heritage into data in the first place. “Digital archives can house so much data that it becomes impossible for scholars to evaluate the archive manually, and organizing such data becomes a paramount challenge”, as some humanities–computer science researchers have stated.²⁰

Accessing digital archives through algorithmic search has, for example, of lately developed into a distinct way of moving beyond the search box. Broadly speaking, a search algorithm is an algorithm—that is a set of procedures and instructions—for finding an item with specified properties among a collection, stored either individually as records in a database or as elements within a search space. The ability to share image data or major cultural data sets at full resolution, however, is crucial for all computational scientists and digital humanities scholars. But with the right properties,

algorithmic software can be applied and used in numerous graphical analysis as well as advanced shape segmentation of digitized heritage. Interestingly, the same software can also be applied to different forms of material, be they illuminations in old paintings or geographical patterns in historical maps. With detailed replicas of objects to be analyzed, shape segmentation algorithms in one particular project (focusing on the study of historical maps) has, for example, been applied to study of medieval manuscripts in another.²¹

The sometimes mentioned research initiative, Digging into Data Challenge, has been one way to tackle these issues. The idea behind the challenge, which hitherto has funded almost a hundred international research teams, has been to address how ‘big data’ has changed the research landscape for the humanities and social sciences. “Now that we have massive databases of materials used by scholars in the humanities and social sciences—ranging from digitized books, newspapers, and music to transactional data like web searches, sensor data or cell phone records”—what forms of computationally-based research methods can be applied? Since the world is becoming increasingly digital, what new techniques will actually be needed to “search, analyze, and understand these everyday materials?” The projects undertaken within the Digging into Data Challenge has, as a consequence, devoted themselves to various forms of ‘big data’ analyzes, often grounded in a digitized cultural heritage. One project has, for example, mined data with ‘criminal intent’ and developed tools and models for comparing, visualizing and analyzing the history of crime using the Old Bailey Online and its extensive court records of almost 200,000 individual trials from 1674 to 1913. Another, “Digging into the Enlightenment” focuses on more than 50,000 digitized 18th-century letters, and analyzes “the degree to which the effects of the Enlightenment can be observed in the letters of people of various occupations.”²²

Another similar research initiative is the so called Cultural Analytics, proposed and undertaken by Lev Manovich at University of California, San Diego. In short, it refers to the use of digital image analysis and visualization for exploring massive visual data sets, and can be seen as a developing methodology within the digital humanities. How to analyze “millions of digitized visual artifacts from the past?”, a description online states. “How do we explore billions of visual born-digital artifacts (both user-generated content and professional media)? How do we research interactive media processes and experiences (evolution of web design, playing a video game)?” To address such challenges, Cultural Analytics has developed methods, techniques and software and

applied these to progressively large data sets. These techniques can, and have been used, within various humanistic disciplines as game studies or media studies, as well as museum exhibitions.²³

For Cultural Analytics—a term naturally linked to Google’s similar offers—the notion of search, however, remains as puzzling as ever. In the before mentioned blog post, Manovich claimed that humanities and media studies researchers today have access to unprecedented amounts of media—“more than they can possibly study, let alone simply watch or even search.” Nevertheless, the basic methods employed “which worked fine when the number of media objects were small—see all images or video, notice patterns, and interpret them—no longer works.” New research models and upgraded ways of seeing are needed, since according to Manovich, standard interfaces for massive digital media collections as, for example, list, gallery, grid, or slide do not allow one to see “contents of a whole collection.” Such interfaces regularly display only a few items at a time, which is an analogue access method that doesn’t allow, or grant a subtler and more sophisticated digital understanding of “the ‘shape’ of overall collection”, nor notice interesting patterns that might emerge.²⁴

If critically examined, projects granted by the Digging into Data Challenge or undertaken within Cultural Analytics could be perceived as somewhat naive in their technological optimism. Contemporary critique of ‘info vis’, for example, often ridicules a similar simple-minded tech positivism, where the notion of ‘more data is better data’ only leads to the paradoxical production of even more (visual) data. Quantitative methods do, after all, have their inherent problems, and even computers don’t replace the need (sometimes) for human interpretation. Yet, what Cultural Analytics in particular has proposed is, actually, way more inclusive cultural histories and analysis of digitized heritage that ideally could take into account “all available cultural objects created in particular cultural area and time period”. A completely digitized history of *all* moving images would, for example, look radically different than today’s canonical film history of artistic masterpieces and commercial blockbusters. Hence, the digitization of massive amounts of cultural artifacts, and the progress in computational tools that can process huge amounts of data, do in fact, make possible a radically new approach to the humanities, not to mention a promise to move beyond simple ‘search’. Mining data, in this respect, also means that humanist scholars no longer have to choose between “data size and data depth.” Rather, they can potentially study “exact trajectories formed by billions of cultural expressions and conversations in

space and time, zooming into particular cultural texts and zooming out to see larger patterns.”²⁵

Conclusion: The Politics of Data

“Think Culture”, runs the subtitle on the heritage portal, Europeana. It’s in many ways a Paneuropean political project with the overall purpose to boost ‘Europeanness’. The portal, however, is promising, even though a giant blank search page still awaits every user when entering it. ‘Search the collections’ isn’t explicitly stated, but inherently users are supposed to follow a standard logic. Then again, “Europeana always connects you to the original source of the material so you can be sure of its authenticity.”²⁶ After all, what’s presented is ‘European Heritage’, not *web n.0* UGC.

At the same time, the notion of simple search has also been put into question, and even been critically examined, as the guiding principle for access to cultural heritage material in general. Digital technology ought, of course, to be used to improve accessibility to Europe’s cultural and scientific heritage in the future. Consequently, Europeana has launched a ThoughtLab where users can explore new initiatives, “participate and have your say, by viewing the demonstration models and sending feedback.” Interestingly one of the models (or projects) presented has the headline, “New ways of searching and/or browsing”. A number of projects are listed, as for example “A Semantic Search Engine For Europeana” that links data together for improved search, or the “Europeana 4D”, an interface that enables comparative visualization of multiple queries and supports data annotated with so called time span data.²⁷ The latter is also related to the project, Europeana Connect, which aims to deliver core components (interfaces for mobile devices, rights licenses, user behaviors etcetera) essential for the development and enhancement of the portal.

If this article has tried to discuss the emergence of new search modalities, the experiments undertaken within Europeana ThoughtLab is but one example where alternative forms of accessibility and novel interface design is trying to move beyond the white search box. Information visualization is, as has been stated, a growing field—and new forms of sophisticated data management another. If Google Image Swirl organizes image search results based on visual and semantic similarities—by way of analyzing pixel values and presenting them in an ‘intuitive exploratory interface’—the Google Public Data Explorer makes large datasets easy to explore, visualize and

communicate. As charts and maps animate over time, changes are “easy to understand”, as Google claims online. Simplicity is, then, key. “You don’t have to be a data expert to navigate between different views, make your own comparisons, and share your findings.”²⁸

Google has, of course, spearheaded such new forms of information retrieval, and tried to go beyond traditional text and hyperlink analysis to unlock information stored in, for example, image pixels. The same goes for the heritage sector—and the co-operation between various national libraries across Europe and Google is an illustrative case in point. Problems with funding digitization activities is, naturally, one reason for such partnerships, but more important for memory institutions is the transfer of know-how. “If our search algorithms can understand the content of images and organize search results accordingly, we can provide users with a more engaging and useful image-search experience.”²⁹ In fact, ‘big data’ approaches are currently being applied to a wide variety of ‘search problems’—all in an effort, perhaps, not to move against search, but definitively towards a more dynamic mode of accessibility to media material and other forms of digital content. From books to maps to the structure of the web itself, ‘the world’s information’ is one “amazing dataset”, as some googlers recently stated in relation to the Google Books Ngram Viewer.³⁰

As Richard Wright, reminds however, there are always politics involved in any representation of data. More data might be better data, yet there are also implicit structures “in digital data, even when, and especially when, that expression takes us far from the realm of computer code.” Then again, the greatest material distance between “human senses and computer code, when compared to the simplest material connections between them, delineates the imaginative possibilities of data visualization”, as Wright densely puts it. According to him this is an area where we currently can “explore the most extreme perspectives that software can create of itself”—its ability to put “cognitive and affective modes of perception into creative tension with data structures and with each other.”³¹ Such affective modes of perception are often as simple as they are convincing, and regularly backed by one commercial interest (or the other). Then again, being human in the digital age constantly means deploying new perceptual functions and modalities, whether affective or not, all in order to cope with the ‘information overload’ paradigm. As this article has shown, one current trend (or strategy) is moving beyond mere ‘search’, all in an effort—and with the expectation—to find new interesting patterns that digital technology can offer.

¹ For a snap shot of Google's first kept search interface using Wayback Machine, see – <http://web.archive.org/web/19981111184551/http://google.com/> (15 September 2011). For an intriguing, as well as graphically mesmerizing presentation of the history of Google's search interface – which makes use of all the available Google front-pages in the Internet Archive – see Richard Rogers *et al.*, “The Demise of the Directory” – <https://wiki.digitalmethods.net/Dmi/DemiseDirectory> and <https://movies.digitalmethods.net/google.html> (15 September 2011).

² Steven Levy, *In the Plex* (New York: Simon & Schuster, 2011), 7.

³ For a discussion, see Geert Lovink & Sabine Niederer (eds.), *Video Vortex Reader: Responses to YouTube* (Amsterdam: Institute of Network Cultures, 2008) – http://networkcultures.org/wpmu/portal/files/2008/10/vv_reader_small.pdf – as well as, Alexander R. Galloway & Eugene Thacker, *The Exploit* (Minneapolis: University of Minnesota Press, 2007).

⁴ Lev Manovich, “Against Search”, blog post at manovich.net, 21 July 2011 – <http://manovich.net/2011/07/22/against-search/> (15 September 2011).

⁵ See, Europeana ThoughtLab – http://www.europeana.eu/portal/thoughtlab_semanticsearching.html (15 September 2011).

⁶ Dion Hinchcliffé, “How social media and big data will unleash what we know” *ZDNet* 12 May 2011 – <http://www.zdnet.com/blog/hinchcliffé/how-social-media-and-big-data-will-unleash-what-we-know/1533> (15 September 2011).

⁷ For a general introduction and discussion on “digital methods”, see Richard Rogers, “The End of the Virtual—Digital Methods” (2009) – http://www.govcom.org/rogers_oratie.pdf – as well as the site Digital Methods Initiative – <https://www.digitalmethods.net/Digitalmethods/WebHome> (15 September 2011).

⁸ See, “BiblioLabs and the British Library Announce British Library 19th Century Historical Collection App for iPad”, British Library press release, 7 June 2011– <http://pressandpolicy.bl.uk/Press-Releases/BiblioLabs-and-the-British-Library-Announce-British-Library-19th-Century-Historical-Collection-App-for-iPad-4f6.aspx> (15 September 2011).

⁹ Alexis Madrigal, “Did the New York Public Library Just Build the Magazine App of the Future?” *The Atlantic* 18 May 2011.

¹⁰ For more information about the *Biblion* app, see – <http://exhibitions.nypl.org/biblion/worldsfair/> (15 September 2011).

¹¹ James Grimmelman, “The Google Dilemma” *New York Law School Review* vol. 53, 2008/09 – http://works.bepress.com/james_grimmelman/19/ (15 September 2011).

¹² Manovich 2011.

¹³ Patricia Cohen, “Digital Keys for Unlocking the Humanities’ Riches” *New York Times* 16 November 2010.

¹⁴ *Our Cultural Commonwealth* ed. Marlo Welshons, American Council of Learned Societies, 2006, 16 – <http://www.acls.org/cyberinfrastructure/ourculturalcommonwealth.pdf> (15 September 2011).

¹⁵ For more information, see – <http://www.ucl.ac.uk/transcribe-bentham/> (15 September 2011).

¹⁶ Melissa Terras, “Present, Not Voting. Digital Humanities in the Panopticon” 10 July 2010 – <http://melissaterras.blogspot.com/2010/07/dh2010-plenary-present-not-voting.html> (15 September 2011).

¹⁷ “Books from 16th and 17th centuries now in full-color view”, Inside Google Books, 17 May 2011 – <http://booksearch.blogspot.com/2011/05/books-from-16th-and-17th-centuries-now.html> (15 September 2011).

¹⁸ Robert Darnton, “Google’s Loss: The Public’s Gain” *New York Review of Books* 28 April 2011.

¹⁹ See the British Library at www.bl.uk, as well as Europeana at – <http://www.europeana.eu/portal/aboutus.html> (15 September 2011).

²⁰ Michael Simeone *et al.*, “Digging into data using new collaborative infrastructures supporting humanities-based computer science research” *First Monday* no. 5, 2011 – <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/3372/2950> (15 September 2011).

²¹ *Ibid.*

²² For information about the Digging into Data Challenge as well as a number of granted projects, see – <http://www.diggingintodata.org/> (15 September 2011).

²³ For a description and introduction to Cultural Analytics, see – <http://lab.softwarestudies.com/2008/09/cultural-analytics.html> (15 September 2011).

²⁴ Manovich 2011.

²⁵ See the description and introduction to Cultural Analytics at – <http://lab.softwarestudies.com/2008/09/cultural-analytics.html> (15 September 2011).

²⁶ See – <http://www.europeana.eu/portal/aboutus.html> (15 September 2011).

²⁷ For a discussion on Europeana ThoughtLab, see – <http://www.europeana.eu/portal/thoughtlab.html> (15 September 2011).

²⁸ See – <http://www.google.com/publicdata/home> (15 September 2011).

²⁹ “Explore Images with Google Image Swirl”, Google Research Blog post, 23 November 2009 – <http://googleresearch.blogspot.com/2009/11/explore-images-with-google-image-swirl.html> (15 September 2011).

³⁰ “Culturomics, Ngrams and new power tools for Science”, Google Research Blog post, 10 August 2011 – <http://googleresearch.blogspot.com/2011/08/culturomics-ngrams-and-new-power-tools.html> (15 September 2011).

³¹ Richard Wright, “Data Visualization”, *Software Studies. A Lexicon* ed. Matthew Fuller (Cambridge, Mass.: MIT Press, 2008), 82.