CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Visualising humanities data
Authors	Day, Shawn
Publication date	2022-11-03
Original Citation	Day, S. (2023) 'Visualising humanities data', in O'Sullivan, J. (ed.) The Bloomsbury Handbook to the Digital Humanities. London: Bloomsbury Publishing, pp. 211-219.
Type of publication	Book chapter
Link to publisher's version	https://www.bloomsbury.com/us/bloomsbury-handbook-to-the-digital-humanities-9781350232129/
Rights	© 2022, Shawn Day. All rights reserved.
Download date	2026-11-16 19:18:08
Item downloaded from	https://hdl.handle.net/10468/13891

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

CHAPTER TWENTY

Visualizing Humanities Data

SHAWN DAY (UNIVERSITY COLLEGE CORK)

Visualization of data is undertaken for a variety of reasons, uses, and purposes in the humanities. Ultimately this forms part of a process of knowledge construction through exploration and discovery. The act of visualizing data as information is both an individual inward pursuit as well as an external performance. Engagement with the viewer/participant and audience raises questions, provokes discussion, and can stimulate activism. Traditionally non-humanistic disciplines have tended to often focus on using data visualization specifically for analysis and definitive substantiation. Until recently, few data visualization tools have been created specifically to fulfill the humanities' unique needs, which has led to adoption and adaptation, often involving conscious or unconscious compromise towards heuristic ends. As a result, these otherwise-engineered tools and methods pose challenges to visualizing humanities data. This chapter explores these challenges and issues to encourage reflection and possibly inspire effective remedy.

Humanistic scholarly practice has a natural aversion to drawing boundaries and putting things into conveniently defined boxes. Non-humanistic disciplines regularly define and construct models as frames of study, which has informed the tools built to visualize data in these spheres. This process implicitly leads to this acceptance of clearly delineated objects of study that follow definable rules. Any sense of inductive reasoning in the humanities pauses before seeking to expound a generalized and abstracted principle. As a result, humanities scholars have sought and devised new and innovative means and coupled them with parallel processes that engage collaboration, crowdsourcing, and empower the public humanities for social change. Discussion of visualizing humanities data is premised on the assumption that this process is unique in many ways from that carried out in other spheres and disciplines. However, this assertion demands critical reflection and consideration of practice.

Visualization of humanities data is an increasingly valuable pursuit, and a "playful iterative approach to quantitative tools," as suggested by Van Zundert, "can provoke new questions and explorations" (Van Zundert 2016). However, the less than critical adoption of the tools, techniques, and methods devised to visualize humanities data (whether developed for specific purpose or adapted from outside) and the potentially compelling but misleading conclusions that might be drawn based on the authority of visual presentation of data pose challenges to scholarly reception (D'Ignazio and Bhargava 2018). This tension demands exploration to raise both awareness and conscious and deliberate activities to mitigate the dangers.

This chapter considers the visualization of humanities data across five axes of tension to expand on the body of scholarship providing critical appreciation:

1. defining versus provoking
2. engaging versus convincing

3. performing versus processing
4. ambiguity versus precision
5. innovating versus demonstrating/adhering.

In *Dear Data*, Giorgia Lupi and Stephanie Posavec (2016) describe a year in which they actively sought to represent their everyday lives through data and creative visualization of this data. The correspondents had only met one another in person twice previously. So, they undertook an experiment to discover how they could get to know one another by exchanging these creative visualizations of each other's lives.

The data visualization process is recounted as playful engagement wherein the correspondents create and share unique and bespoke visualizations of their everyday lives. This particularly engaging book sparks two critical observations. First, human activity (personal or collective) can and is being increasingly metricized. Second, this data (and the seriously playful part) can be represented in various visual ways that bring the data back to life and allow for the rhythms and cycles of ordinary existence to be engagingly explored and shared for deeper understanding.

Lupi and Posavec highlight how humanities scholars today engage in play with data that often results in forms of visualization that create new forms of *representation* and similarly connect this with *performance*.

Early reflections on the particularities of humanities engagement with data, such as Drucker's seminal reflection on the constructivist origins of captured "data," have led to a critical realization of how humanities and social science scholars must make explicit the connection to the sources and the representational nature of data (Drucker 2011). These questions around data as subjective and "given" (Galloway 2011) in contrast with information as "something which is taken" have critical implications for what we represent visually, let alone how we choose to do so. Lavin has recently defended the term data as being "far more rich etymologically" than Drucker credited it but reaffirmed nonetheless the critical engagement with data that marks a unique humanities approach (Lavin 2021).

Similarly, data can be repeatedly re-represented through consequent processes and undergo a reductionist transformation (much as imagery does through the xerographic process). However, it can also be distorted and enhanced through collation and conjunction with related data—simplification and legibility, making the underlying data less transparent to the viewer (Drucker 2011).

So, what makes the visualization of *humanities* data unique?

One lately emerging observation is the unique value that humanities scholars can bring to the scientific process. Where data becomes sacrosanct and its existence as *capta* is overlooked, the perceived authority of numbers and data—once captured—is rarely reflected upon. Instead, humanities scholars by practice are encouraged to bring a natural skepticism or curiosity to the origins as a corrective, to step back and consider sources and acquisition to identify misuse through bad data practice.

Rising to Drucker's call for "humanists to think differently about the graphical expressions in use in digital environments" (2010), the subsequent decade witnessed an appreciation of information visualization as a tool not merely for new ways to conceptualize data through graphical means but also in the social affordances of the process of visualization. Visualization is the shiny, distracting toy that entices but distracts.

In 2010 Joanna Drucker challenged the discipline to "a critical frame for understanding visualisation as a primary mode of knowledge production" as she presciently pointed out, at the time, "critical understanding of visual knowledge production remains oddly underdeveloped." So, where do we find ourselves today? How have we met this challenge in the humanities? Jennifer Edmond posed a challenging counterpoint to the ordering of knowledge, asserting that much like messy desks, our modes of reading in the humanities are where the serendipity happens (2018). We need to maintain multiple layers of engagement to make the mental leaps that result in synthesis and knowledge assembly. So how do we find the use of visualization in the humanities has evolved? How has self-reflection informed our engagement with the multivariate forms of data that we work with in the humanities today? So, what does the hackneyed phrase "appealing directly to visual cortex" really mean? Bypassing language centers and avoiding overt translation and processing, plugging directly in cerebral appreciation suggests a rawness, but simultaneously a blunt but genuine reality—unprocessed, unrefined, and yet with an appealing authenticity.

DEFINING VERSUS PROVOKING

The allegation that data visualization provides a means for STEM practitioners to pack conclusions into tidy boxes contrasts with a distinctly different way in which similar processes are employed in the humanities to inspire new questions.

As a reflection on humanities practice, Edmond explores the practice of reading in academic practice and argues that distraction and the "noise" of the deluge of data is an intrinsic part of how we make connections and are inspired in knowledge creation (2018). She further demonstrates how this practice has changed and is changing by the impact of dealing with increasing quantities of data in our everyday life, whether on our digital media or in the way in which traditional media has shifted to data-driven (and visually presented) narrative.

Likewise, this ability to playfully engage with the data through imaginative visualization has called upon the artistic and creative manner in scientific practice that is not limited to the humanities scholar; William Allen points to this playful engagement as a form of brokerage practiced throughout society (2018). He highlights the nature of the relationship that develops through engagement between the creator, the disseminator, and the recipient and demonstrates how this is a reciprocal process of essentially play. Anecdotally, the author recalls a surreal night on the town in Cambridge, MA, where he met a researcher at MIT. Although a practicing fine artist, he was also undertaking DNA sequencing in a lab at MIT, where his imagination was seen as crucial to the scientific process. However, where this creativity or a unique appreciation of the human condition may be seen as unique to humanities scholarship (it is not), embracing what is perceived as a scientifically rigorous means of engaging with data is also integral and thoughtful collaboration.

We need to play in creation but also play in engagement. Edward Tufte affirmed the role of visualization of data to "create engagement" in his seminal *The Visual Display of Quantitative Information* (Tufte 2001). The use of visualization to analyze and present data in purely scientific pursuit is used to substantiate and reinforce findings. The humanities, however, adopt a seeming counterpose. Visualization of humanities data (as illustrated by Lupi and Posavec) invites questioning and discovery, provoking engagement and seeking to inspire questions instead of substantiating findings. This is indeed a far different approach and intent.

Likewise, there remain aspects of humanities data and how it is conceptualized for visualization. Humanities data is often inherently messier than that in scientific pursuits due to the human and the condition being considered. Representation of ambiguity, for example, is of increasing focus in many spatial humanities circles. These have led to conscientious approaches to develop and explore new techniques to visualize particular to the humanities addressing gray areas and recognize the need not to imply greater precision or certainty to the data through the way it is portrayed.

ENGAGING VERSUS CONVINCING

However, visualizing data in the humanities is not merely about discovery and exploration. Still, it perhaps uniquely has an implicit responsibility to stimulate engagement and interaction with the broader community (Fitzpatrick 2019) to activate phenomenological reflection on data. Thus, humanities information visualization, although a form of representation for realizing, identifying, or extracting knowledge, is equally employed for stimulating critical thought to enable action, participation, etc.

Although Helen Kennedy et al. (2016) assert that visualization primarily attempts to make data represented objective, this is contrary to the underlying appreciation in the humanities that data is unambiguously subjective, and that visualization seeks to affirm that fact and provoke subsequent reflection. Objectivity remains a STEM determination, an attempt to convince that discussion has been exhaustively pursued and concluded beyond a reasonable doubt. For the humanities scholar, the visualization is meant to suggest, connect, provoke, and propose in the hope of instigating further discussion, living up to the adage that *when the conversation stops, the visualization fails.*

The ability to carry out effective data visualization is integral to engaging with different layers of abstraction. Simply presenting all the available data often hinders the process and overwhelms the viewer (Galloway 2011). Abstracting allows for a broad visual perspective of a large dataset (a God's eye view) but utilizing emerging technological affordances allows for data visualization that permits examination in detail where desired and allows for critical engagement with the underlying data (Tufte 2006). Emerging forms of visualization such as multi-layered tree-maps (Bederson et al. 2002), sunbursts (Feitelson 2004), or packed circles (Wang et al. 2006) are particularly engineered to afford this engaged mode of visual interaction.

This trend towards the use of hierarchical data visualization as a visual interface to collections has been growingly apparent in the realm of cultural heritage where collection artifacts (grounded in well-constructed metadata) have engendered new forms of visual browsing techniques. For example, collection browsing platforms, such as the 100 Archive[1] or Vikus Viewer,[2] offer "generous interfaces" and allow for leveraging metadata schema for enhanced exploration and discovery.

The affordance of visualizing humanities data densely, as a means of cultural observation, has inspired the questions: What if our visual representations showed all the available datapoints? Could this allow us a more informed perception of previously invisible social phenomena? And if so, what new forms of visual representation would this demand? (Manovich 2011). All of which suggest the pronounced power of data visualization to allow individuals to bring their own experiences and contexts to richer, deeper, and larger unlocked datasets which were previously unfathomable in tabular or textual forms.

These affordances further amplify and empower the call by Kathleen Fitzpatrick for the active involvement of scholars and, in particular, those in the humanities to ensure its survival and that of humankind through the reinvention of the university's role (2019). The ability to communicate research and share datasets more effectively through visualization can amplify access to scholarly output by wider society. It reaffirms the role of the humanities scholar (and not humanities alone) outside of the environs of academia.

However, this, in turn, demands a digital criticality that allows the public sphere to be drawn into a renewed engagement eased through the responsible and informed use of data visualization. Visualization of humanities data provides the power and the opportunity for engagement, and it harnesses the traditional humanities performative practice.

PERFORMING VERSUS PROCESSING

In the humanities, novel or innovative approaches are expected practice and innovation itself is often a demanded part of the process. Arguably, there are fewer convenient frameworks to be applied—or in fact, demanded in the humanities—but a unique approach is readily acknowledged and accepted. Although there are defined methods and a general epistemic approach, the beauty of humanities practice, especially concerning the visualization of data, is in the performance and is often in the way we journey (Ellenberg 2021). Thus, the process of visualizing data and the aspects of the process undertaken are not just critical to its outcome but lend an experience throughout the performance of visualizing humanities data (Kräutli and Davis 2016).

There has been a perceptible increase in scholarship on both how performance has a role in what we would typically see as a purely scientific pursuit and necessary critical reflection on the nature of the process and how the original data were compiled. The intrinsic link between the representations and the represented is a vital aspect of visualization and one which Lev Manovich raises by characterizing tool and method development as a conversation. His work challenges a rote "process" of visualization of data. It suggests the process as a way of thinking and of "talking out" the process of deliberating and making manifest our thought processes, "a challenge to traditional forms of cultural analytics" (Manovich 2009). The development of tools and methods to visualize humanities data demonstrates how algorithms and interfaces are crafted through code—the means of approaching the problem is a form of the language of representation and analysis. Manovich's Software Studies lab reimagines and reinvents the visual interface to access and analyze massive amounts of human behavioral data. These interactive visualizations permit analysis and questions such as "What happens when you can see an entire body of artistic output within a single frame?" or "Can we explore cross-cultural behaviors across time and space through social media?" (SelfieExploratory[3]). Big questions that are left open-ended and to the viewer to answer through their own exploratory experience.

In a manner, this recognition that the journey (process) has much to contribute to the outcomes of an exercise of visualization is not unique or new but connects to the practice of the performative in the humanities and an openness and willingness in many humanities disciplines to the open-ended and the serendipitous as a recognized form of scholarship. Access to, and appreciation of, foundational artistic aspects highlight the importance of the process, whether in notational representation as a form of visualization of data, or in finding means to develop visual interfaces to navigate the immensity of social media data to explore human behavior.

AMBIGUITY VERSUS PRECISION

Perceived end use determines the varying levels and degrees of precision required in visualization (as well as the underlying data itself). The term "perceived" is particularly noted here as the publisher certainly cannot imagine all potential ways a visualization may be interpreted. The need for a means to represent ambiguity or fuzziness in visualization to more honestly convey uncertain or ambiguous data has long been realized. However, the responsibility incumbent on the creator, author, or publisher of the visualization to represent this uncertainty has only been more recently demanded. Geospatial visualization has struggled with this challenge (Plewe 2002; Grossner and Meeks 2014). This need for "fuzziness" in geographic information systems used to visualize both spatial and temporal data is challenged by digital representation. The binary and fixed nature of static forms of Cartesian display is inadequate to represent disputed boundaries, imperfect memory, or contested interpretation. Over time we have devised means to attempt to represent ambiguity, displaying contradictory claims, using transparency as a proxy for precision, or even utilizing new visualization models, such as augmented reality, allowing the user to experience diverse "realities." However, in many of these cases, the digital nature of the data itself lends itself to a determined precision, precision that may not, in fact, represent the true nature of humanity or human phenomena. Likewise, how precision—or ambiguity—is represented has massive implications for interpreting and using the represented (or mis represented) data (Grossner 2020). One of the most significant challenges confronting most consumers of visualized data is realizing that it is not definitive, that it is based on fuzzy data and needs to be recognized as such when attempting to draw conclusions. The increasing prevalence of what is referred to as "deterministic construal error" must be addressed but requires a digital competency that remains unrealized mainly (Joslyn and Savelli 2021) not just within academia but throughout society.

Digital methods are increasingly embracing ambiguity as being genuinely representative and thus a crucial aspect of being, but this is not an aspect of traditional or analog visualization practice. Although some forms of visualizations such as the whisker diagram—used to represent fluctuating stock prices over a particular period—have allowed for a degree of explicit brackets of representation, even this demands a border/fixed frame of reference.

The digital challenges humanities scholars, but in pursuit of bringing honesty and transparency to emerging powerful tools, it demands unique and innovative means to allow for varied precisions and representations that avoid implying structure or definition where none exists.

INNOVATING VERSUS DEMONSTRATING/ADHERING

There is a perceptible impetus in the humanities subtly demanding that the creation of new knowledge requires innovation in process, where non-humanistic practice demands greater adherence to specific methods. Although ultimately stemming from data supporting replicability as part of the scientific process, in the humanities, replication has not typically been seen as an essential part of practice, and conclusions are most often open-ended. The performative aspect mentioned previously often demands innovation in process or at least in application—counter to traditional non-humanistic—and a habit amongst those visualizing humanities data.

All the challenges we have identified lead us to consider where innovation exists in the humanities today and how other spheres are innovating to deal with emerging challenges. The scientific

method proposes a rigid methodology that can be employed to systematically explore phenomena through data to reach a replicable and determined solution to a hypothesis. The challenge in the humanities is that there is a less rigid application of the methodology and an embracing of exploration and discovery—often without a predetermined hypothesis. Unlike STEM practice, there is not a demand for a conclusive and definitive outcome. A conclusive definition demands adherence and more rigid adherence to demonstrated practice than to innovative engagement (D'Ignazio and Bhargava 2018). Ambiguity is embraced, represented, and made transparent—if well visualized.

When the process becomes performance and developing new forms of visualization becomes an end in itself, the culture of reflection marries with the culture of innovation.

CONCLUDING

Data from the outset has been and remains contentious, not just in the humanities. Before even considering visualizing for whatever intent, the nature of the object under observation and analysis has raised questions about bias, significance, completeness, form, etc. Issues raised by Drucker and others have sparked critical discussion about the underlying layer of the visualization process, particularly in the humanities. The result of this discussion which we do not engage with in depth in this chapter, is that we often (usually?) start with a flawed object, and then we visualize.

In visualizing data, we then make decisions around methods, aesthetics, intent that may distort further but certainly cannot repair what was not already there. If the data is flawed in any manifold ways mentioned earlier, we only move away from mimetic truth. The challenge then becomes whether we can stimulate discussion around the represented data, to illustrate or at least make transparent its faults and thereby use it to bring greater understanding if not just recognition that we need better data.

Visualization of data in any discipline serves various purposes, and the humanities are not distinct in this. However, this chapter has sought to illustrate that there are unique uses and processes in the humanities that distinguish practice and demand ongoing reflection, especially considering the ever-growing authority and pervasive power of visualization in modern society.

In the simplest structural form of practice, information visualization can be bifurcated into analysis and presentation. Within these spheres, it serves additional and varied functions. But ultimately, after the analysis and the results, how information is visualized is based on a scholarly propensity to share. We all feel this necessity to translate and share knowledge to give it a broader and longer life and find an impactful purpose. This simply reaffirms the mutually beneficial collaboration between diverse approaches to visualizing data in the humanities that invoke performance, invite engagement and provoke rather than palliate.

NOTES

1. See The 100 Archive, https://www.100archive.com.
2. See The Vikus Viewer – Urban Complexity Lab – University of Applied Sciences Potsdam, vikusviewer.fh-potsdam.de.
3. Software Studies Laboratory, http://selfiecity.net/selfiexploratory/.

REFERENCES

Allen, William L. 2018. "Visual Brokerage: Communicating Data and Research through Visualisation." *Public Understanding of Science* 27 (8): 906–22. https://doi.org/10/gf7wd4.

Bederson, Benjamin B., Ben Shneiderman, and Martin Wattenberg. 2002. "Ordered and Quantum Treemaps: Making Effective Use of 2D Space to Display Hierarchies." *ACM Transactions on Graphics* 21 (4): 833–54.

Brüggemann, Viktoria, Mark-Jan Bludau and Marian Dörk. 2020. "The Fold: Rethinking Interactivity in Data Visualization." *Digital Humanities Quarterly* 14 (3).

D'Ignazio, Catherine and Rahul Bhargava. 2018. "Creative Data Literacy: A Constructionist Approach to Teaching Information Visualization." *Digital Humanities Quarterly* 12 (4).

Drucker, Johanna. 2010. Graphesis: Visual Knowledge Production and Representation. http://scholar.google.com/scholar_url?url=https://journals.tdl.org/paj/index.php/paj/article/download/4/50/0&hl=en&sa=X&ei=Ny5JYNH7McbYmQHZl554&scisig=AAGBfm22wpsUJ0DUArb5ogFJBUQAXS-OJQ&nossl=1&oi=scholarr.

Drucker, Johanna. 2011. "Humanities Approaches to Graphical Display." *Digital Humanities Quarterly* 5 (1).

Edmond, Jennifer. 2018. "How Scholars Read Now: When the Signal Is the Noise." *Digital Humanities Quarterly* 12 (1).

Ellenberg, Jordan. 2021. *Shape: The Hidden Geometry of Absolutely Everything*. London: Penguin.

Feitelson, Dror G. 2004. "Comparing Partitions with Spie Charts." *Occasional Paper, School of Computer Science and Engineering, The Hebrew School of Jerusalem*. https://www.cs.huji.ac.il/~feit/papers/Spie03TR.pdf.

Fitzpatrick, Kathleen. 2019. *Generous Thinking: A Radical Approach to Saving the University*. Baltimore, MD: Johns Hopkins University Press.

Galloway, Alexander. 2011. "Are Some Things Unrepresentable?" *Theory, Culture & Society* 28 (7–8): 85–102. https://doi.org/10/fxjbmv.

Grossner, Karl and Elijah Meeks. 2014. "Topotime: Representing Historical Temporality." Conference Paper DH, Lausanne. July. https://www.researchgate.net/publication/272827034_Topotime_Representing_historical_temporality.

Grossner, Karl. 2020. "Towards Geospatial Humanities: Reflections from Two Panels." *International Journal of Humanities and Arts Computing* 14 (1–2): 6–26.

Joslyn, Susan and Sonia Savelli. 2021. "Visualising Uncertainty for Non-Expert End Users: The Challenge of the Deterministic Construal Error." *Computer Science* 27 (January). https://doi.org/10.3389/fcomp.2020.590232.

Kennedy, Helen, Rosemary Lucy Hill, Giorgia Aiello, and William Allen. 2016. "The Work that Visualisation Conventions do." *Information, Communication & Society* 19 (6): 715–35. https://doi.org/10.1080/1369118X.2016.1153126.

Kräutli, Florian and Stephen Boyd Davis. 2016. "Digital Humanities Research through Design." *DH Early Career Conference 2016 – Mapping the Scope & Reach of the Digital Humanities*. King's College London, May 20.

Lavin, Matthew. 2021. "Why Digital Humanists Should Emphasise Situated Data over Capta." *Digital Humanities Quarterly* 15 (2).

Lupi, Giorgia and Stephanie Posavec. 2016. *Dear Data: A Friendship in 52 Weeks of Postcards*. Hudson, NY: Princeton Architectural Press.

Manovich, Lev. 2009. "Cultural Analytics: Visualizing Cultural Patterns in the Era of 'More Media'." *DOMUS*. http://manovich.net/content/04-projects/063-cultural-analytics-visualizing-cultural-patterns/60_article_2009.pdf.

Manovich, Lev. 2011. "What is Visualisation?" *Visual Studies* 26 (1): 36–49. https://doi.org/10/dg3k7t.

Manovich, Lev. 2015. "Data Science and Digital Art History." *International Journal for Digital Art History* 1. https://journals.ub.uni-heidelberg.de/index.php/dah/article/view/21631.

McCosker, Anthony and Rowan Wilken. 2014. "Rethinking 'Big Data' as Visual Knowledge: The Sublime and the Diagrammatic in Data Visualisation." *Visual Studies* 29 (2): 155–64. https://doi.org/10/gh8w2z.

Plewe, Brandon. 2002. "The Nature of Uncertainty in Historical Geographic Information." *Transactions in GIS* 6: 431–56. doi: 10.1111/1467-9671.00121.

Software Studies. 2014. The Selfieexploratory. http://selfiecity.net/selfiexploratory/.

Tufte, Edward R. 2001. *The Visual Display of Quantitative Information*, 2nd edn. Cheshire, CT: Graphics Press.

Tufte, Edward R. 2006. *Beautiful Evidence*. Cheshire, CT: Graphics Press.

van Zundert, Joris. 2016. "Screwmeneutics and Hermenumericals: The Computationality of Hermeneutics." In *A New Companion to Digital Humanities*, edited by S. Schreibman, R. Siemens, and J. Unsworth, 331–47. New York, NY: John Wiley and Sons.

Wang, Weixen, Hui Wang, Dai Guozhong, and Hongan Wang. 2006. "Visualisation of Large Hierarchical Data by Circle Packing." *Conference Proceedings of the 2006 Conference on Human Factors in Computing Systems*, CHI 2006, Montréal, Québec, Canada, April 22–27. doi: 10.1145/1124772.1124851.

Whitelaw, Michael. 2015. "Generous Interfaces for Digital Cultural Collections." *Digital Humanities Quarterly* 9(1).