

## 9. Approaching data visualizations as interfaces: An empirical demonstration of how data are imag(in)ed

*Daniela van Geenen and Maranke Wieringa*

### Abstract

This chapter points out data visualization's double role as explorative and communicative means in humanities research. We draw from science and technology studies looking at the mediation process at stake: the interaction between visualization tool and researcher. To emphasize this mediation process and expose the various decisions at its heart we introduce the term 'data interface'. We highlight *how* visualizations function as data interfaces and visualization practices allow for *interfacing with* data biographing a network graph's 'life'. Using the lens of the 'data interface' underscores that a particular (network) visualization provides just one perspective on the data. Moreover, we examine *if* and *how* the used data interfaces encourage scholars to critically position their investigative work, during research processes and communication.

**Keywords:** Data interface; Critical positioning; Mediation; STS; Visual network analysis

### Introduction

In the introduction of *Science in Action*, science and technology studies scholar Bruno Latour (1987) illustrates how particular scientific findings and technological developments led to the three-dimensional model of DNA with which we are familiar today. This introduction is the prelude to Latour's call to study the production of knowledge 'in the making', instead of merely

focusing on outcomes (1987, p. 4). In this chapter, we respond to Latour's call to bring the epistemic process, and particularly, the ways in which data are imag(in)ed, to the foreground of the research communication. Our approach 'stages' the 'cultural life' of a data visualization in scholarly research. That is to say, we write a biographical account that depicts the visualization's making and distribution. This approach alludes to the double role of data visualization: first, visualization as an activity employed during the research process to get a perspective on data by the application of 'exploratory data analysis' (EDA) (Tukey, 1977). Second, data visualizations as images used as representations and research results, which are publicly communicated (e.g. Lynch & Woolgar, 1990; Coopmans et al., 2014).

We point out data visualization's double role by looking at the interaction and negotiation between the software tool used for visualizing the data and the researcher using this tool. The notion of the 'data interface' is introduced to emphasize this mediation process. By biographing the 'life' of a single network visualization, we pinpoint its role as a data interface. The network graph was created in a humanities research project, which investigated the Dutch-speaking Twittersphere as communication infrastructure (van Geenen et al., 2016). Using the case of this specific visualization, created with the network visualization tool Gephi (Bastian et al., 2009), we underscore how the various steps and decisions in the construction and the subsequent circulation of a graph play a role in defining it.

## Defining 'data interfaces'

We look at data visualizations by approaching them as interfaces. Interfaces 'are the point of juncture between different bodies, hardware, software, users, and what they connect to or are part of' (Cramer & Fuller, 2008, p. 150). As such, interfaces function as mediators between different entities in situations in which 'users are not simply the audience, but also the actors' (Chun, 2011, p. 65). Like the graphical user interfaces (GUIs) which are familiar to us, (the process of) data visualization mediates between the data and its beholder. This mediation is especially emphasized when graphical representations of data are offered as, and have been derived from an interaction with, a GUI. Software scholar Wendy Chun (2011) aptly noted that GUIs, which offer tangible entry points to engage with abstract information, should be understood as 'programmed visions'. This notion pinpoints the non-neutral, preprogrammed quality of the graphical (re) presentations interfaces present to the user.

Moreover, ‘programmed vision’ implies that certain visions of the developers, implemented into the software programs, are reified by the use of these tools. In the mediation processes featured by these tools, the actual execution of the underlying code stays invisible, thus obscuring the developers’ choices (Chun, 2011). The exploration presented in this chapter highlights the mediation process, in which the preprogrammed and, therefore, inscriptive quality of software plays an important role. We show that this ‘situatedness’ of the research methods, in combination with the scholars’ academic, cultural, and social background (Haraway, 1988), is vital to knowledge production. The notion of the ‘data interface’, then, emphasizes the sense-making process of visualization performed by both the researchers and the future ‘readers’, who are faced with the data visualization as research outcome.

We are not the first to consider the interfacing aspect of data visualization: practitioners like Citraro and Rees (2015) have made a good case for a complementary line of argument. Stephen Few (2014) and Gray et al. (2016) likewise note the mediative character of data visualizations. In contrast and addition to these previous approaches, our aim is to demonstrate *how* data visualization figures as data interface in diverse situations, in scholarly practice and (public) research communication.

## Taking account of the ‘life’ of data visualizations

Gephi, the software tool we used for our exploration, is a popular open-source software program for mapping, manipulating, and analysing all kinds of network data (Bastian et al., 2009). The software tool was designed to enable social and cultural scholars with little technical expertise to encounter complex relational data at the level of the GUI (Heymann, 2010). Consequently, it is often used in humanities and social science research. Its designers present Gephi as a tool for ‘Visual Network Analysis’ (Heymann, 2010; Venturini et al., 2015), thereby placing the emphasis on interfacing with and ‘reading’ network visualization. In this contribution we wish to sketch two reading positions with regard to network visualizations: that of the researcher engaging in exploratory data analysis (EDA), and that of an audience member to whom research results are offered via the (scientific) image of a network diagram.

EDA was coined by statistician John Tukey (1977), and is often one of the initial stages of a research project. In this stage, researchers are striving to grasp the examined phenomenon and to comprehend the data on which they are working (O’Neil & Schutt, 2014, pp. 34–36). EDA uses plots, summary statistics, and most applicable to our contribution, graphs (O’Neil & Schutt,

2014, p. 35). EDA using Gephi features Social Network Analysis (SNA) (Bastian et al., 2009; Jacomy et al., 2014), which is a branch of the social sciences that builds on mathematical principles of graph theory to chart interpersonal relations and examine social structures (Marin & Wellman, 2011). Insights in SNA are derived from the graph, as the positions of nodes—for example, individuals—are dependent on their connectedness and thereby the positions of all other nodes (Marin & Wellman, 2011).

Thus, EDA mobilizes particular forms of knowledge and methodological principles. The rise of the application of software tools in humanities and social science research has prompted some scholars to reflect on their effect on the research process and outcomes through the ways in which ‘our digital helpers are full of “theory” and “judgement” already’ (Rieder & Röhle, 2012, p. 70). We account for the way in which our ‘digital helper’ frames the research process by outlining the interaction with Gephi and stressing the relevant steps we take to make sense of the data *and* the tool.

In relation to visualizations’ communicative capacity, it is not so much the exploratory process of gaining insights into the data that is relevant. Rather, the network graph provides a very specific kind of ‘interface’ to the data, displaying one of many possible perspectives on this research material. In other words, it is here that the visualization’s—unnoticed—rhetoric power and, therefore, the question of understandability come to the fore (e.g. Haraway, 1988; Kennedy, Hill, Aiello, & Allen, 2016; Latour, 1986). In this contribution we discuss *how* a visualization functions differently at various stages of its life by focusing on specific aspects of its rhetoric power and pointing out how it requires particular forms of ‘reader’ engagement.

## The life of a network visualization

Before a network visualization—or any information visualization—is ‘born’, data need to be selected, extracted, cleaned of irregularities in data formatting, and filtered on specified parameters. It is only after the data have been prepared for analysis that we usually start to visualize them. This, however, does not mean that the visualization stage is the final stage of the research. We will discuss how visualizations can also feed back into one’s analysis, and thereby, become a particular kind of interface for working with the underlying data. To do so, we will biograph a network visualization with which both authors are familiar: a network visualization displaying day-to-day communication practices (@replies) in the Dutch-speaking Twittersphere (van Geenen et al., 2016).

## Staging the mediation process

In the beginning of our network graph's life, we departed from tabular information extracted from Twitter's application programming interfaces (APIs) (see van Geenen et al., 2016 for detailed information on the corpus collection). The collected data sample contains more than 3.5 million Dutch tweets sent between 4 and 12 September 2016 (van Geenen et al., 2016). After cleaning the data (i.e. fixing formatting errors and dealing with missing information due to the partially black-boxed data extraction from Twitter), we started preparing them for an exploration in Gephi. As we were interested in communication between users, we filtered out solely replies (i.e. tweets that start with @username). Simultaneously, we added the usernames of accounts these replies addressed as an additional column to the spreadsheet. In doing so, we were accommodating the use of Gephi in our analysis, since the tool requires two types of data points in order to visualize relations as the foundation for the network graph: a source and a target. The following sections will concentrate on the mediation process in Gephi, exploring the data visually, on the one hand, and preparing the network graph as communicable visualization, on the other.

### Gephi's focus on sociality

Gephi's analytical strength resides in its layout algorithms (Bastian et al., 2009). The 'ForceAtlas 2' layout algorithm was specifically developed for use in the Gephi application software and is optimized for handling large sets of relational data (Jacomy et al., 2014, pp. 5-11). The application of ForceAtlas 2 is stimulated by Gephi's design and stimulated by the Gephi core team (see e.g. van Geenen, 2018, for a more elaborate discussion of this matter). This technical specification makes it suitable for the processing and exploration of our dataset, which consists of 224,305 nodes (accounts), connected by 499,485 edges (809,871 sent replies; in case of double connections these were merged to weighted and thus thickened edges). Put in motion in 'Overview', one of Gephi's three tabs, ForceAtlas 2 causes a gradually perceivable spatialization of the graph. Next to 'Overview', Gephi features 'Data Laboratory' (i.e. allowing for inspection of the tabular data) and 'Preview' (i.e. allowing preparation and export of the static network graph). 'Overview' plays a vital part in knowledge production in Gephi, in the data processing and graph spatialization. The ForceAtlas 2 spatialization is force-directed. Thus, connections (replies, in our case) attract nodes (accounts) whereas nodes themselves repulse each other (van Geenen, 2018, pp. 2-3).

This simulation clusters the graph based on the number of connections nodes possess (degree), a clustering principle termed ‘modularity’ (van Geenen, 2018, p. 2). The application of modularity can be understood as a ‘distant reading’ strategy that features the lens of sociality. This strategy helps structuring the EDA approach to a large dataset that comprises social interactions. The research project at stake studied Twitter as everyday communication infrastructure (van Geenen et al., 2016). In order to identify this infrastructure, we used Gephi’s ‘Modularity Class’ community detection algorithm (set to resolution 0.5 to identify also smaller communities), which classifies nodes based on shared connections (Blondel et al., 2008). Starting from a single node, the calculation process ‘snowballs’ through the entire graph and measures with which cluster each node has the most connections, and based on this, generates node metadata. Subsequently, we coloured and ‘partitioned’ the nodes based on the communities inferred by the algorithm.

While we are aware of the flaws of representing modularity in such a fashion, we used this strategy with a quantitative orientation as an initial exploration to follow up with a ‘close reading’ of these clusters. Modularity does not express, for instance, whether a particular node is strongly or loosely affiliated with a particular cluster, which erases the nuances we touched upon. We used modularity to initiate the qualitative encoding of the encountered clusters, and simultaneously, question the validity of these ‘inferred data publics’ (de Lange, 2017) based on the nodes’ connectedness. In that we performed a close reading of both the research material *and* the data interface.

### A close reading of visual network analysis in Gephi

According to Gephi developers, the ForceAtlas 2 algorithm provides ‘transparency’ in offering a continuous, manipulable simulation of the graph spatialization process (e.g. Jacomy et al., 2014, p. 2). There is a feeling of directness when working with the program, especially when the algorithm is ‘running’, set into operation by the push of one button. As with large graphs such as our communication network, it takes time to render the spatialization and reach a point at which the node positioning is moderately stable. Yet, when algorithm properties are tweaked during this process, through the settings panel, one sees the network visualization’s instantaneous response (see Figures 9.2 and 9.3). While Figure 9.1 shows the ‘raw’ graph, Figure 9.2 displays the graph after running the algorithm, adapting and experimenting with the ‘Scaling’ (i.e. the adaptable graph size that takes the node positioning into account) and the ‘Gravity’ (i.e. the simulated forces

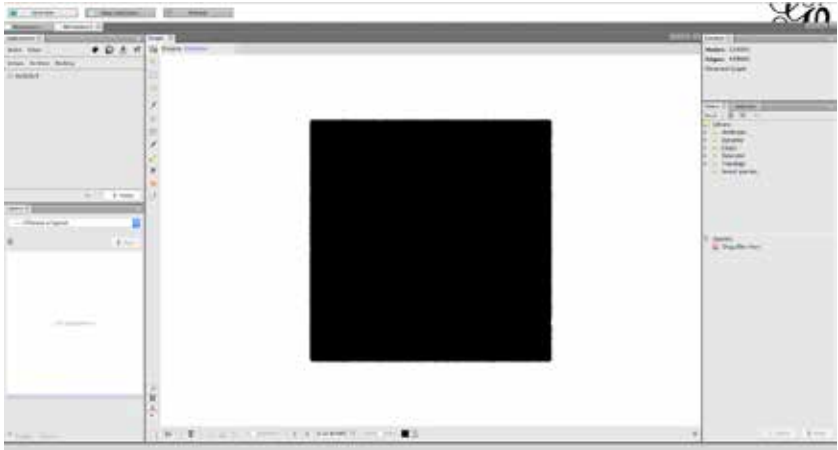


Figure 9.1. ‘Raw’ version of the network graph in the ‘Overview’ after the data import into Gephi. Created by D. van Geenen using Gephi.

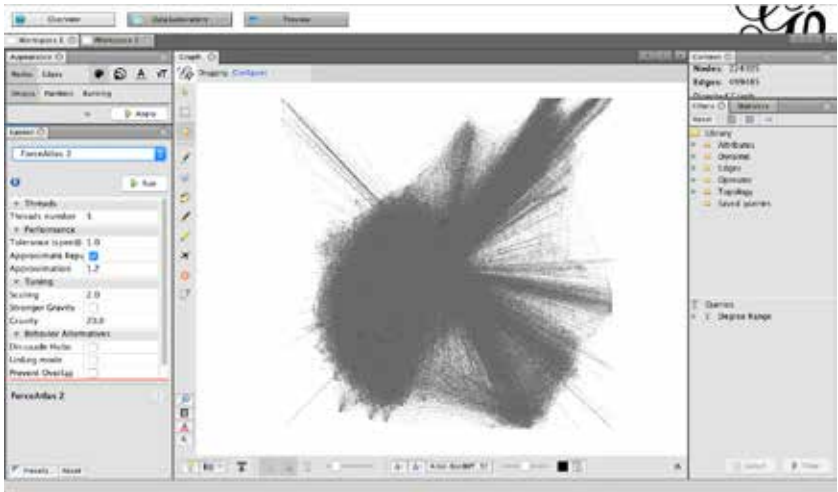


Figure 9.2. Spatialized graph after the application of ForceAtlas 2 (Scaling: 2.0, Gravity: 20.0; node size based on degree). Created by D. van Geenen using Gephi.

attracting nodes to the centre) of the graph. Thus, through Gephi’s software affordances, the designed action possibilities the tool offers (e.g. Gaver, 1991), the visualization itself becomes an interface to the data, offering (the promise of) ‘direct manipulation’ (Shneiderman, 1982; for detailed analysis of Gephi’s affordances, see van Geenen, 2018).

Direct manipulation can be described as the ‘representation of the object of interest, rapid incremental reversible actions and physical action instead of complex syntax’ (Shneiderman, 1982, p. 237). It can be understood as an

immediate, visual feedback on a user's given action. For network visualization in Gephi, most of these characteristics of 'direct manipulation' apply. An exception are conveniently reversible actions, as Gephi does not offer an 'undo button' (cf. van Geenen, 2018). In tweaking the algorithm, though, settings can be 'reversed' by means of changing properties such as scaling back to the previous value, which results in a similar node positioning. (Since the software program presents a graph *simulation*, exact node positions can slightly differ.)

Apart from tweaking the running layout algorithm, applied filters are another example that have a kind of 'live' effect on the appearance of the network graph (e.g. Bastian et al., 2009). Based on the detected modularity clusters, we started filtering out all extremely small communities, which appeared to be unconnected from the main graph. Moreover, as a strategic focus in the preparation of the close reading of the graph, we decided to delete all nodes with less than ten connections (degree). In other words, we chose to concentrate on the most active accounts, which had sent or received more than ten replies (Figure 9.3). To sum up, we initially approached the algorithmic processing of the data with an EDA strategy: 'playing around' with the layout algorithm's settings and filters through direct manipulation in order to come to a first legible spatialization. Whereas the visualization in Figure 9.1 is not helpful in providing insights into the data, the graph spatialization assists in 'reading' the data (see Figure 9.2). This spurred subsequent tweaking, research questions, and explorations, resulting in Figure 9.3 as the 'final' graph.

### Situating Twitter publics beyond modularity

For the purpose of situating the identified communities we close read the profile information of a sample of accounts per cluster to define and classify these communities. We combined these observations with the knowledge that we had gained through our long-term engagement with the Dutch Twittersphere. The graph presented us with the 'usual suspects' in communication research on Twitter: a dense cluster of highly connected, politically interested professionals such as politicians, media organizations, and journalists. However, modularity also opened the way for a new perspective on the data: we saw users coming together around particular occupations (e.g. foresters, or people involved in teaching and education), topics (e.g. public debate on sustainability), or interests (e.g. the Dutch theme park *De Efteling*). Some of these communities were beyond our expectations, for instance the gaming and vlogging communities. The visualization,



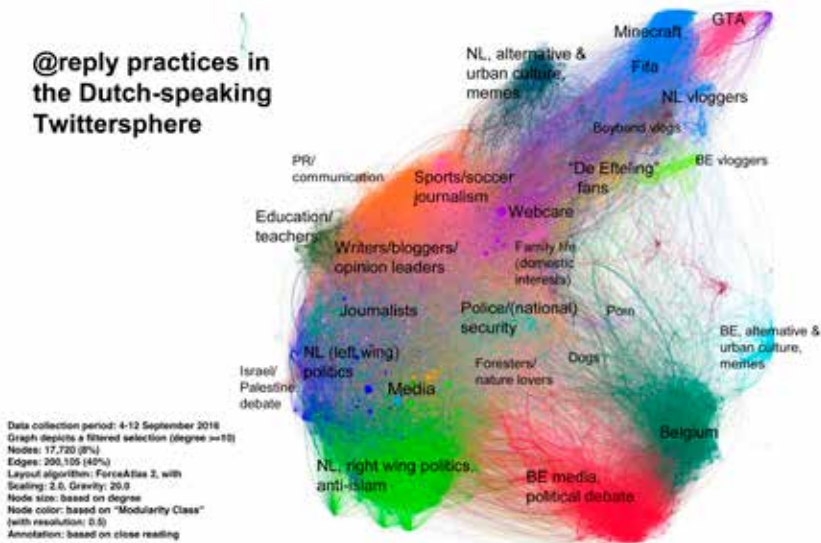


Figure 9.3. Exported, spatialized, filtered, and annotated graph. Created by D. van Geenen and M. Wieringa using Gephi and Photoshop in preparation of conference presentations.

then, became an interface to the data: it added to the perception of the research material by transforming tabular data into a palpable structure and manipulable form.

The network graph above provides more clues about the data such as different kinds of detectable media practices, that is the diverse ways in which users made use of Twitter's specifications such as the @reply functionality. As we studied reply practices, webcare accounts 'skewed' our sample. These are accounts of diverse (commercial) organizations that provide customers with the opportunity to address their concerns about products and services. The webcare account of the Dutch Railways (@ns\_online) was in fact the most connected account in our graph (with 4,968 edges compared to an average degree of 3.6 edges for the whole graph), followed by @postnl and @kpnwebcare. We discussed whether we should exclude such webcare accounts, which due to their interaction with a diversity of other accounts function as central connectors in the reply network. Since the spatialization (solely) builds on the 'social hierarchy' of degree, it does not discriminate between the different media practices. Eventually, we decided to include all the different media practices, using the network visualization as a point of departure for further research.

Other research sparked by this graph included an investigation of the media practices of Dutch-speaking politically interested communities and the 'locality' of Twitter publics (see van Geenen et al., 2016). The first

study originated from the observation that particular politics groups are well-represented in the communication infrastructure, such as accounts that interacted with the official account of right-wing politician Geert Wilders, the first politician in the list of highly connected accounts. Here we also observed that many of the tweets sent in this cluster surrounding @geertwilderspvv were concerned with a local incident in the Dutch city of Almere. It deepened our interest in the dynamics between the national and local spheres of public debate, and the role of local engagement in everyday communication on Twitter.

### **Between exploration and communication**

We understand the function of mapping the data in the shape of an evolving network graph, and in this interfacing with the data, as ‘augmenting human intellect’, to borrow from interface design pioneer Douglas Engelbart (1962). In the context of EDA, then, network visualization helps scholars to come to a ‘degree of comprehension in a situation that was previously too complex’ to fathom (Engelbart, 1962, p. 1). The practice of visualization can be said to further scholarly thought, by making sensible what otherwise remains an overload of tabular information (e.g. Gray et al., 2016). As we have demonstrated, network graphs can figure as useful data interfaces.

However, we also need to be aware of what they do and do not show, and for what reasons. That is, we need to take the time to reflect *in which way* network visualization imposes shape on our research. Donna Haraway (1988) argues for the need to account for the positioning of the researcher, when she considers ‘situated knowledge’. It has been argued elsewhere that, since network analysis and visualization thrive on software, the tools researchers use and become familiar with should be considered part of their positioning (van Geenen, 2018). In light of Haraway’s observation, the biography of the network visualization we are sketching accounts for our research practice, since it depicts our engagement with relational data, based on our backgrounds and, thus, the (preliminary) knowledge we draw upon. With this account, then, we reveal our own ‘critical positioning’ and the ‘partial perspective’ (Haraway, 1988, pp. 585–586) we have on the research material as scholars. By doing so, we show that graphs do not exist in a void, but come into being through a complex series of interactions, based on certain preset conditions, of which the decisive moments should be made available to the public.

## The cultural circulation of a network visualization

At some point in the research process we have to start formalizing our findings and translate them into something communicable to an audience. We used Gephi's 'Preview' tab to tweak the readability of the visualization. Furthermore, the tab's interface affords exporting the graph as static image (i.e. as PNG, PDF, or SVG file): a 'screenshot' of the sense-making process. Here, the network visualization—literally—moves from a state of mutability to a form of 'immutability'. While EDA draws on the graph's mutability, for example, by means of tweaking the layout algorithm's settings, the preparation of the findings for communication purposes results in a single static representation of the graph, which can be used in all kinds of (scholarly) publications. It becomes an 'immutable mobile' (Latour, 1986, pp. 7-13), which, as we discuss below, affects the public's possibility to engage critically with the presented research findings.

After annotating the exported network diagram in Photoshop, this image has led a particularly 'eventful' life: the visualization was presented at academic conferences (van Geenen et al., 2016) and has been featured on television for a broader audience (Boeschoten, 2017). Furthermore, we touched upon its complexities at the Impakt Festival, an annual, popular scientific festival around new media (van Geenen & Wieringa, 2017). Below, we will discuss how this 'screenshot' functions in two of these contexts: the academic conferences and the Impakt Festival.

### Academic conferences: Scientific mediation versus public research communication

In presenting our research at several conferences, we were faced with a familiar dilemma articulated by numerous STS scholars: in which ways should we make use of graphical representations, which are expected to be the 'objective product' of a systematic knowledge production process (e.g. Haraway, 1988; Latour, 1986)? Moreover, such static images can easily be shared due to scholars' access to media platforms that invite sharing visual information such as Twitter. During and after our presentations we found the prepared network diagram circulating on Twitter, some versions with more comprehensive annotations on its making process than others. Due to their compressed nature, such 'screenshots', we argue, do not live up to the dynamic data interfaces that helped to imagine the data. This observation encouraged us to think of forms of 'methodological reflexivity' (Rieder & Röhle, 2012, p. 80) in the research communication that could do justice

to the complex mediation process from which the network visualization originated. For instance, we developed an internal distribution policy: a code of conduct for the contextual information, which should be featured on slides and included in papers or non-academic articles (e.g. Figure 9.3).

Furthermore, for the purpose of catering to the traceability of, and stimulating reflection on, the meaning-making process of researchers interfacing with the data, we built a 'fieldnotes' plug-in for Gephi (Wieringa et al., forthcoming). It provides a comprehensive time-stamped version of both a text file of the applied settings in Gephi and a network graph file.

### **Impakt: Tackling data interfaces in interaction with the public**

In our contribution to the Impakt Festival we demonstrated how single static images do not do justice to the complexities of the network graph (van Geenen & Wieringa, 2017), if only because graphs are nearly illegible if they are comprised of a vast amount of nodes. During our Impakt presentation, we elaborated and reflected on the network graph's making: rendering the research process visible, allowing the public critical engagement with the visualization.

Our talk addressed the diversity of reply practices the visualization represents and argued that forms of procedural mapping and interactive engagement should be on the agenda for (critical) data studies approaches. We exemplified the mediation process in Gephi to the public in a video that showed how ForceAtlas 2, in interaction with the researcher, handles the data. In this we made an effort to confront the idea that vision, especially that of an expert viewer—such as ourselves—working with algorithmic, standardized visualization tools, will automatically lead to absolute objectivity (e.g. contributions to Coopmans et al., 2014). This problem was aptly expressed in Haraway's notion of the 'god trick' (1988, p. 589), a phenomenon that is amplified through the 'programmed visions' our data interfaces present us with. Showing and explaining the video to the public, we *mobilized*—literally and figuratively speaking—the network graph to demonstrate the importance of reflection on such visualization practices. Concluding, we made an effort to demonstrate how data can be imagined and imaged, and how the resulting visualization is streaked with particular norms, conventions, and rhetoric (Haraway, 1988; Kennedy et al., 2016).

To summarize, we used the network graph as an illustration stating that data interfaces feature the potentiality to 'augment human intellect'. However, this augmentative capacity depends on their affordances to

facilitate a continuous assessment of the principles applied to make sense of the data. As such, we vouch for ‘account-ability’ (Garfinkel, 1967 as cited in Eriksén, 2002) implemented into tool design processes. We ended our talk on data interfaces at Impakt Festival advocating for the need to think of design strategies that provide the opportunity to access data interfaces in (actual) interaction. During our talk we positioned our work critically. Our objective in writing this article is to further stimulate research in which data interfaces are approached critically, and that questions how modes of ‘tool criticism’ could be built into our data interfaces.

## Conclusion

In conceptualizing the process of data visualization as a form of interfacing, we have added new perspectives on data visualization. The focus on data interfaces stimulates an understanding of the process of data visualization as mediation, and the resulting image as one of many possible interfaces to the data. Using the vehicle of the biography of a data visualization, we highlighted how visualization practices allow for *interfacing with* data, and exposed the choices and selections at the heart of this process. We emphasized data visualizations’ constructedness, and thus their role as results of a specific knowledge production process. We examined in which ways the data interfaces we use are capable of encouraging the scholars to critically position their work during the research process. By doing so, we argue that researchers should practise tool criticism (cf. van Es, Wieringa, & Schäfer, 2018; van Geenen, 2018). Moreover, we strive to stimulate them to provide the audience of their research outcomes with the possibility of a critical engagement with their network graphs.

## References

- Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An open source software for exploring and manipulating networks. *Proceedings of the Third International ICWSM Conference*, 361-2. Retrieved from <https://aaai.org/ocs/index.php/ICWSM/09/paper/view/154>
- Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), P10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008>

- Boeschoten, T. (2017, December 4). Hoe schrijven boze witte mannen onze geschiedenis? [Video file]. Retrieved from <https://www.universiteitvannederland.nl/college/hoe-schrijven-boze-witte-mannen-onze-geschiedenis/>
- Chun, W. H. K. (2011). *Programmed visions: Software and memory*. Cambridge/London: MIT Press.
- Citraro, D., & Rees, K. (2015). Epiphanies through interactions with data. In: D. Bihan (Ed.), *New challenges for data design*. (pp. 175-185). London: Springer-Verlag.
- Coopmans, C., Vertesi, J., Lynch, M., & Woolgar, S. (2014). *Representation in scientific representation revisited*. Cambridge/London: MIT Press.
- Cramer, F., & Fuller, M. (2008). Interface. In: M. Fuller (Ed.), *Software studies: A lexicon*. (pp. 149-152). Cambridge/London: MIT Press.
- de Lange, M. (2017, March & April). *Data publics in the smart city*. Paper presented at the Data Publics conference, Lancaster University, UK. Retrieved from <https://themobilecity.nl/2017/11/24/my-keynote-talk-at-data-publics-conference-lancaster-university-march-31-april-1-2-2017/>
- Engelbart, D. (1962). *Augmenting human intellect: A conceptual framework*. Summary Report AFOSR-3233. Menlo Park, CA: Stanford Research Institute.
- Eriksén, S. (2002). Designing for accountability. *Proceedings of the Second Nordic Conference on Human-Computer Interaction*, 177-186. <http://doi.org/10.1145/572020.572041>
- Few, S. (2014). Data visualization for human perception. In: Interaction Design Foundation (Ed.), *The encyclopedia of human-computer interaction* (2nd ed.). Aarhus: Interaction Design Foundation.
- Garfinkel, H. (1967). *Studies in ethnomethodology*. Englewood Cliffs: Prentice-Hall.
- Gaver, W. W. (1991). Technology affordances. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 79-84. <https://doi.org/10.1145/108844.108856>
- Gray, J., Bounegru, L., Milan, S., & Ciuccarelli, P. (2016). Ways of seeing data: Toward a critical literacy for data visualizations as research objects and research devices. In: S. Kubitschko & A. Kaun (Eds.), *Innovative methods in media and communication research*. (pp. 227-251). Cham: Palgrave Macmillan.
- Haraway, D. J. (1988). Situated knowledges: The science question in feminism and the privilege of partial perspective. *Feminist Studies*, 14(3), 575-599.
- Heymann, S. (2010, February 1). Gephi initiator interview: How 'semiotics matter' [Blog post]. Retrieved from <https://gephi.wordpress.com/2010/02/01/gephi-initiator-interview-how-semiotics-matter/>
- Jacomy, M., Venturini, T., Heymann, S., & Bastian, M. (2014). ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PLoS ONE*, 9(6), 1-12. <https://doi.org/10.1371/journal.pone.0098679>

- Kennedy, H., Hill, R. L., Aiello, G., & Allen, W. (2016). The work that visualisation conventions do. *Information, Communication and Society*, 19(6), 715-735. <https://doi.org/10.1080/1369118X.2016.1153126>
- Latour, B. (1986). Visualization and cognition: Drawing things together. *Knowledge and Society*, 6, 1-40.
- Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. Cambridge, MA: Harvard University Press.
- Lynch, M., & Woolgar, S. (1990). *Representation in scientific practice*. Cambridge/London: MIT Press.
- Marin, A., & Wellman, B. (2011). Social network analysis: An introduction. In: P. Carrington and J. Scott (Eds.), *The sage handbook of social network analysis*. (pp. 11-25). London: Sage.
- O'Neil, C., & Schutt, R. (2014). *Doing data science: Straight talk from the frontline*. Sebastopol: O'Reilly.
- Rieder, B., & Röhle, T. (2012). Digital methods: Five challenges. In: D. M. Berry (Ed.), *Understanding digital humanities*. (pp. 67-84). London: Palgrave Macmillan.
- Shneiderman, B. (1982). The future of interactive systems and the emergence of direct manipulation. *Behaviour and Information Technology* 1(3), 237-56. <https://doi.org/10.1080/01449298208914450>
- Tukey, J. W. (1977). *Exploratory data analysis*. Reading: Addison Wesley.
- van Es, K., Wieringa, M., & Schäfer, M. T. (2018). Tool criticism: From digital methods to digital methodology. *Proceedings of the 2<sup>nd</sup> International Conference on Web Studies*, 24-27. <https://doi.org/10.1145/3240431.3240436>
- van Geenen, D. (2018). *The role of software tools in critical data studies practices: Studying the affordances of Gephi as a sociotechnical system* (Master's thesis, Utrecht University, Netherlands). Retrieved from <https://dspace.library.uu.nl/handle/1874/367489>
- van Geenen, D., Schäfer, M. T., Boeschoten, T., Hekman, E., Bakker, P., & Moons, J. (2016, October). *Mining one week of Twitter: Mapping networked publics in the Dutch Twittersphere*. Paper presented at The 17th Annual Conference of the Association of Internet Researchers, Berlin, Germany.
- van Geenen, D., & Wieringa, M. (2017, October 26). *Data visualization as interface* [Video file]. Retrieved from <https://youtu.be/QPGcJkHqpxc?t=42m17s>
- Venturini, T., Jacomy, M., & Pereira, D. (2015). *Visual Network Analysis*. Working Paper. Retrieved August 15, 2016 from [https://www.tommasoventurini.it/wp/wp-content/uploads/2014/08/Venturini-Jacomy\\_Visual-Network-Analysis\\_WorkingPaper.pdf](https://www.tommasoventurini.it/wp/wp-content/uploads/2014/08/Venturini-Jacomy_Visual-Network-Analysis_WorkingPaper.pdf)
- Wieringa, M., van Geenen, D., van Es, K., & van Nuss, J. (forthcoming). The Field Notes Plugin: Making network visualization in Gephi accountable. In: A. Daly, K. Devitt and M. Mann (Eds.), *Good Data*. Amsterdam: Institute of Network Cultures.

## About the authors

**Daniela van Geenen** is a lecturer and researcher in data journalistic practices and data visualization at the University of Applied Sciences Utrecht and an affiliated researcher at the Datafied Society research platform. She investigates the scholarly ethos that working with digital methods calls for and studies the design of accountable software tools.

**Maranke Wieringa** is a PhD candidate at Utrecht University, where she is part of the Datafied Society research platform and teaches various courses on (scholarly) data analysis. Her academic interests lie at the intersection of software and data. For her dissertation, Maranke investigates algorithmic accountability in Dutch municipalities.