Participatory Data Design: Acting in a digital world

Jensen, Torben Elgaard; Birkbak, Andreas; Madsen, Anders Koed; Munk, Anders Kristian

Published in:
Making & Doing

Publication date:
2021

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Jensen, T. E., Birkbak, A., Madsen, A. K., & Munk, A. K. (2021). Participatory Data Design: Acting in a digital world. In G. Downey, & T. Zuiderent-Jerak (Eds.), *Making & Doing: Activating STS through Knowledge Expression and Travel* (pp. 117-136). MIT Press.

Making & Doing

Activating STS through Knowledge Expression and Travel

The title-level DOI for this work is:

doi:10.7551/mitpress/11310.001.0001

5

PARTICIPATORY DATA DESIGN

Acting in a Digital World

Torben Elgaard Jensen, Andreas Birkbak, Anders Koed Madsen, and Anders Kristian Munk

In a series of recent projects, the Techno-Anthropology Lab (TANTlab) and its collaborators have developed an approach to STS making & doing, which we now call *participatory data design* (PDD).[1] PDD projects are often organized as so-called data sprints, three- to five-day intensive workshops where STS researchers collaborate with stakeholders in a particular field to collect and generate data that is relevant to a particular problem—and together create a useful digital device or object (Munk, Meunier, and Venturini 2019). Thinking about such collaborations as a form of participatory design raises the question of whether and how PDD experiments can create more equitable and locally beneficial forms of data use by engaging local actors in the work of finding relevant data and of building something meaningful out of it. But the engagement of local actors also comes with a degree of unpredictability. In our experience, these kinds of participatory projects are different every time and often involve a feeling of loss of control and ambivalence about the role of digital data. This chapter therefore does not offer a finished conceptual framework but rather a demo in the shape of a number of stories about our projects, their recurrent features, and some of the dilemmas and tensions that arise when making and doing STS with PDD.

One of the central tensions that runs through the chapter is related to the contemporary imperative to do something digital in academia—and in society more generally. There is currently a strong interest in digital tools across the social sciences and humanities (Vertesi and Ribes 2019), which is something that TANTlab benefits from when generating institutional support and setting up collaborations with external partners. What is most interesting about the current imperative to go digital, however, is perhaps that it can be an opportunity to blur unhelpful dichotomies between qualitative and quantitative (big data) methods, between description and intervention, between online and offline, and between technical insiders and outsiders. The aim of this demo is to show how we try to work across such oppositions in

our concrete makings and doings in TANTlab while also registering and attending to how these tensions constantly reappear in new disguises. Our aim is to ponder the consequences of having data as a boundary object between STS research and external stakeholders. Which role does data play in this context, and what kinds of procedures and goals can set the direction for a PDD process? In the following we give a short situation report from a particular moment in a recent project in which we and our collaborators found ourselves in the midst of the promises and difficulties of doing PDD.

John seemed to be out of his depth.[2] He had been arranging citizen hearings about new technologies for several years. He knew all about putting together the key information material for these hearings: a carefully worded leaflet given to the citizens with a balanced collection of expert views on a particular subject matter. He also knew that he was on solid ground institutionally speaking: his employer, the Danish Board of Technology, was an internationally recognized pioneer of technology assessment and citizen engagement. So basically, the task of organizing another citizen hearing on a technoscientific issue—in this case the dilemmas related to pandemics—should not cause too much commotion.

This time, however, the Danish Board of Technology had engaged with new odd bedfellows: a group of researchers from the TANTlab at Aalborg University who had promoted the idea that material gathered from social media such as Twitter could be used to identify the burning issues that should be brought before the citizen hearing. John and his colleagues from the board had quickly embraced the idea of doing something with Twitter data. They were well aware that more and more public debate takes place on social media platforms, and they were interested in ways to relate this new phenomenon to their existing methods, although also somewhat hesitant because of what they perceived as a lack of conventional methodological checks and balances in digital methods (Birkbak, Madsen, and Munk 2020). The group from the board and from TANTlab had met a few times before, and now they were engaging in a weeklong joint workshop to explore in earnest how large quantities of social media data might play into the work of the Danish Board of Technology.

When preparing a leaflet for a citizen hearing, it might seem a small step to take soundings in Twitter space rather than merely calling on established experts in the field. But as John and his colleagues soon realized, the new approach entailed more than just gathering information from a different source. In the normal process, John would hear or read every single word that the called-upon experts would deliver. In this case, the researchers from TANTlab turned up with a database of four hundred thousand tweets related to the Zika virus! With this massive amount of material, the process of identifying the issues had to rely on different tools and different types of choices. Rather than speaking to trusted experts and asking directly for their views on possible key themes, John and the other workshop participants had to navigate through a data-overwhelmed and thus rather disorienting situation.

The digital tools provided by the TANTlab researchers allowed them to manipulate the large data set in various ways. The details of what was done on the computers of the TANT lab researchers was not entirely accessible or understandable to John and his colleagues. The outcome, however, was more tangible. At various points in time, the researchers printed out large, poster-size maps, displaying hundreds of words distributed in a number of hairball-like clusters, indicating that particular Twitter hashtags frequently occurred together in the material. Everybody gathered around these discursive maps, trying to make sense of the clusters. Whenever an interesting cluster was located on the map, the TANTlab researchers encouraged the participants to invent a tentative thematic label for the cluster and to read through some of the underlying tweets that had given rise to this particular constellation of words. The material would often appear excessive and with a lot of "noise" in the form of words that didn't suggest anything meaningful about pandemic-related issues. Repeatedly, the TANTlab researchers asked questions about how to filter the data: Would it be okay to look only at tweets with two or more hashtags? Would it be okay to look only at hashtag combinations with three or more distinct users? Such filtering decisions made it possible to print out new and potentially less "noisy" maps, which could then be examined.

The iterative process of visualizing, filtering, labeling, and collectively discussing went on for several days. It gradually became clear that the digital methods would not be replacing John's skills in identifying relevant topics for the citizen information leaflet. John's skills and intuitions just had to be incorporated in another way. At one point, the filtering process had returned a limited number of tweets that seemed to cover key aspects of the total material. At this time the workshop participants invented what they half-jokingly called the *John score*. Leaning on John's expertise, they asked John to read through the tweets and rate their relevance for the information leaflet. Did they point to a question or a dilemma that was important for the citizen hearing to address?

The tweets that were evaluated positively by John were incorporated into the information leaflet alongside the expert testimonies. In this way, a space and a role were created for new data opportunities within the normal operating procedures of the Danish Board of Technology. However, this space was not evident from the beginning. For instance, it was created by experimenting with different ways of moving between big data analyses of tweet patterns and close reading by John, who used his intuition as an expert in citizen engagement to make these analyses meaningful for the context in which they were to be used. Such experimental iterations are indicative of the emerging STS practice of PDD. In this chapter, we draw on our experience from several PDD projects and provide suggestions as to how STS researchers can work with external partners around data. We start by examining some of the developments that enable our current experiments with PDD. Then we delve into three key practices in PDD projects: datafication, flexible visualization,

and materialization. Finally, we discuss some of the challenges and consequences of pursuing PDD as a particular kind of engagement experiment.

ENABLING CIRCUMSTANCES

Since 2015, we have experimented with the approach we now call PDD. In the following we describe some of the situated circumstances that enable and configure our work.

TECHNICAL DEVELOPMENTS ENABLING A DIGITAL STS

When the participants from the Danish Board of Technology were exploring the network of co-occurring Twitter hashtags about the Zika outbreak they were engaging in a practice that has now become commonplace but was difficult to imagine only fifteen years ago. It owes its existence to a convergence of digital developments that we can broadly label as digital methods (e.g., Birkbak and Munk 2017; Marres 2015; Rogers 2013).

The first of these developments is the emergence of social media, or what is sometimes known as the Web 2.0, which is characterized by a shift from static information sharing to dynamic, user-generated content on platforms like Twitter or Facebook (O'Reilly 2007). Social media platforms allow us to gather large amounts of structured data on how individual users interact with issues in real time (Lomborg and Bechmann 2014). In the case we described earlier, Twitter allowed us to download data on how users from all over the world were interacting with the hashtag #zika over a period of a month. It allowed us to track what questions they were raising in relation to Zika, which links they were sharing, what arguments they were pushing, and what other users they would typically associate with in doing so. Twitter has become one of the platforms through which a public can express its concerns around an issue, and this offers opportunities for research.

The second development is what we could call the popularization of computational methods for tracking and analyzing digital traces. It is of course possible to simply watch the Zika discussion unfold through the Twitter interface one tweet at a time. But the availability of tools like the Twitter Capture and Analysis Toolkit (Borra and Rieder 2014) for harvesting tweets systematically and the Gephi visual network analysis platform (Bastian, Heymann, and Jacomy 2009) has made it possible to easily conduct exploratory data analysis on large data sets (in this case almost half a million tweets).

METHODOLOGICAL DISCUSSIONS AND PROJECT PRACTICES IN STS AND MEDIA STUDIES

In the 2010s, academic fields such as media studies and (digital) STS have appropriated the new technical opportunities to study what Noortje Marres would call

issue publics (Marres 2005) and what Bruno Latour would call controversies (Latour 2004). In close collaboration with software developers and data visualization experts, STS researchers have contributed to the further development and specialization of digital tools, and in conjunction with this, media and STS scholars have developed an increasingly sophisticated discussion about methodological challenges related to finding, harvesting, visualizing, and interpreting patterns of digital traces. These methodological discussions include reflection on how to think about the relationship between online and offline data (Munk 2013; Rogers 2009); how to bridge the qualitative-quantitative divide (Blok and Pedersen 2014; Munk 2019); how to take into account that each type of social media comes with a specific genre that conveys and configures particular types of messages (Burgess and Matamoros-Fernández 2016; Madsen 2013; Marres 2015); how to handle facts, values, and incommensurable positions in knowledge controversies (Birkbak 2013; Venturini 2010); and how to use data visualizations exploratively to pose new questions (Munk and Jensen 2014; Munk, Madsen, and Jacomy 2019; Venturini, Ricci, et al. 2015).

In parallel with the evolving methodological discussions mentioned earlier, STS scholars have learned some hard-won lessons about the difficulties of making data visualizations that are meaningful and useful to external parties. Early projects (e.g., MACOSPOL) were based on the implicit idea that sufficiently good visualizations would be picked up by the general public as easy-to-use tools for understanding and navigating public knowledge controversies such as discussions about vaccines or climate adaptations (Munk, Meunier, and Venturini 2019). Later projects (e.g., EMAPS) realized the necessity of inviting external stakeholders early and directly into the *engine room* when selecting and analyzing data (Venturini, Meunier, et al. 2015). For this reason, the data sprint format was developed where knowledgeable stakeholders (issue experts) were invited to give presentations of their views of the key challenges of their particular field and where the stakeholders were also involved in the subsequent data analysis and assessment of the data projects that were created through the workshops (Munk, Meunier, and Venturini 2019).

TAKING INSPIRATION FROM THE SCANDINAVIAN TRADITION OF PARTICIPATORY DESIGN

A final enabling circumstance that deserves specific mention is the tradition for user orientation and involvement in the fields of design and IT-development (Birkbak, Bornakke, and Papazu 2017; Callon, Lascoumes, and Barthe 2009; Munk, Meunier, and Venturini 2019), not least found in the Scandinavian tradition of participatory design.[3] Participatory design was originally developed in the 1980s in an effort by trade unions to gain influence on how workers' conditions were changed by the introduction of new production technologies, which the unions feared would lead to deskilling of jobs. The trade unions soon realized that resistance to new technologies

tout court was an impossible strategy, but that valuable influence could be gained by attending to the many different ways that specific technologies could be configured. In principle, it would therefore be possible to negotiate a version of technology that improved rather than diminished the quality of work life. Today, we could raise the same issue in relation to the exponential rise of data and computational power: How can we negotiate a version of data practices that improves rather than diminishes life?

A second key lesson from the 1980s is the importance of the practical tools of collaboration. In the participatory design movement, it was quickly realized that the prospect of developing progressive technologies often hinged on intimate knowledge of previous work processes and conditions. Designers and implementers often lacked this knowledge, but under the right circumstances it could be brought to the table by workers, who could represent future users. This in turn raised methodological questions: What would be the best way of evoking the workers' sometimes tacit knowledge about work processes, and what would be the best way of staging the collaboration and negotiation between designers and future users? In the decades that followed, people in the participatory design movement worked on these methodological questions and developed a considerable repertoire of practical tools and approaches, such as future workshops, design games, mock-ups, and joint prototyping (Kensing and Blomberg 1998).

Our PDD projects are premised on the idea that the development of positive and meaningful uses of data will benefit substantially from direct involvement of the stakeholders, who are the potential future users. We thus share the collaborative strategy of the participatory design tradition as well as a keen interest in any kind of procedures and devices that can help articulate users' knowledge and open designers' decisions for negotiation. This raises questions such as whether the flexible visualizations that we discuss later could be considered a digitized version of the low-fidelity prototypes that characterize participatory design or whether the addition of data changes something more substantial in the participatory situation.

THREE KEY PRACTICES IN PDD

The varieties of enabling circumstances are parts of the background for the actual doing of the projects, which we turn to now. We draw on our experiences from five PDD projects carried out between 2015 and 2018.[4] But rather than describing every single project, we describe three key practices that were recurrent in our attempts to establish meaningful interactions between us and the participants in these projects: *datafication*, *flexible visualization*, and *materialization*.

DATAFICATION

A PDD project is always about a particular problem: How can we use discussions on Twitter to improve the information material for a citizen hearing? How could Aalborg municipality gain sensitivity to public debates on Facebook while implementing a new school reform? How could the Royal Danish Theatre understand its visitors in new ways by looking at patterns of event attendance? When approaching these or other problems we do not simply import large quantities of digital data; we begin by actively engaging in *datafication*. Datafication is the process of gathering a substantial amount of digital data and relating it to a particular problem.[5] When we first meet with possible partners in a PDD project, we ask whether digital data might already exist that relates to the problem. We might ask also if people currently engage in activities that leave digital traces about the problem. Or we might ask if it is possible to actively get people to produce digital material.

The opportunities for gathering data are rapidly increasing as digital data and devices become more pervasive parts of social life, work life, political life, research, economic transactions, and all other sectors of society (Madsen 2015). In specific PDD projects, the datafication is a matter of seizing the opportunities at hand and finding the most relevant and accessible types of data. As a consequence, the types of data we have worked with are quite diverse. In a PDD project about obesity research we harvested a large number of scientific publications from scientific journal databases as well as a large number of food-related Instagram posts; in a project on school reform we and the municipality invited a vast crowd of stakeholders to participate in several months of debate on a specific Facebook page. In a third project, with the Royal Danish Theatre, we traced the activities of a large number of users across multiple Facebook pages.

Deciding on the relevant data source is the beginning of datafication. Other challenges soon follow. Digital data is usually abundant, which means that participants in PDD projects must decide to focus on a selection of the material, such as data generated within a particular time frame or by certain users. Another set of options that open up and need to be negotiated are questions about metadata and cascade effects. To explain this, it is useful to make a distinction between digitization and datafication (Mayer-Schönberger and Cukier 2013). As an example of digitization, we might think of a digital photograph of an old and fragile manuscript written on parchment. The viewer of the digital photograph can do no more or no less than a reader of the original document. The only thing gained is the ability to make cheap copies and to read without the risk of ruining the fragile original. Datafication, by contrast, is taking digitization an important step further. If, for instance, the photograph is scanned with optical character recognition technology and transformed into digital text, the reader will be able to perform a range of additional operations. She can do simple tasks such as a word count or identify the most frequent words in the document. She may also use her digital tools to make comparative analyses

of the frequencies of words in different parts of the text or identity key phrases that often occur together. Or she may deploy digital tools based on machine learning to automatically detect the themes talked about in the text or the mood of its author. The possibility of performing all these operations indicates that datafication is quite different from simple digitization. Datafication allows us to deploy algorithms that add new layers of information (metadata) to the original material. Through this, it becomes possible to perform a cascade of operations that explores new combinations and correlations within the data set. In this way, datafication opens a series of data manipulation opportunities. A key activity of any PDD project is to discover and explore some of these opportunities and discuss if and how they may be of value to the stakeholders and the fields they are a part of.

The process of datafication often stimulates queries, which we might call ontological: What is really going on in this field, and what kinds of data might be relevant to relate to it? What are the essential parts of the phenomenon that we should attempt to articulate with or find in this data? Datafication also tends to stimulate discussions about the normative commitments of the participants: What kinds of analyses and visualizations could, or should, we attempt with this material? These are dilemmas that arise early in a PDD process, and even though they may generate tensions between the participants, these dilemmas make it necessary to be explicit about choices in ways that raise interesting discussions about the role of digital technology in society.

In several projects, we have seen that the value of datafication lies in stakeholders and issue experts discussing *at length* how to operationalize their concerns with respect to a specific data set. For instance, in the project with the Royal Danish Theatre there was an extensive discussion about the notion of authenticity. The participants had no doubt that authenticity was the hallmark of the royal theatre as opposed to other cultural enterprises, such as movie theaters. But when the participants were challenged to operationalize authenticity with respect to a large Facebook data set—when they had to datafy the problem—they found themselves in the midst of a thought-provoking discussion. Would it be possible to identify topics that were more authentic than others? Would ballet be a more authentic topic than football? Would an exclusive event be more authentic than a more popular event? Or was *authentic* a quality that users might indicate by using particular words? All these discussions brought the participants' knowledge and assumptions to the fore as they debated how to operationalize authenticity.

The lesson we draw from this is that PDD projects should avoid the temptation to define the key problems and concepts in advance. PDD should be approached as an opportunity to discuss and define the key problems with the issue experts and make a joint effort to operationalize these ideas in and with specific data sets. The practice of operationalization and datafication carries opportunities to "make our ideas clear" (Peirce 1955, 23). The cost of staying with the trouble in this way is that participants

(including the organizers) will often experience a high level of uncertainty or even frustration because no one can guarantee that the PDD workshop will yield anything meaningful.

In sum, datafication is not a safe road to knowing everything about a field on the basis of massive amounts of data. Instead, datafication can be described as a riskier path that allows the participants to explore difficult ontological and normative questions and to open up possibilities for further development of data. We see this as a mode of pragmatic inquiry that involves redefining problems through empirical experiments (Munk, Madsen, and Jacomy 2019).

FLEXIBLE VISUALIZATION

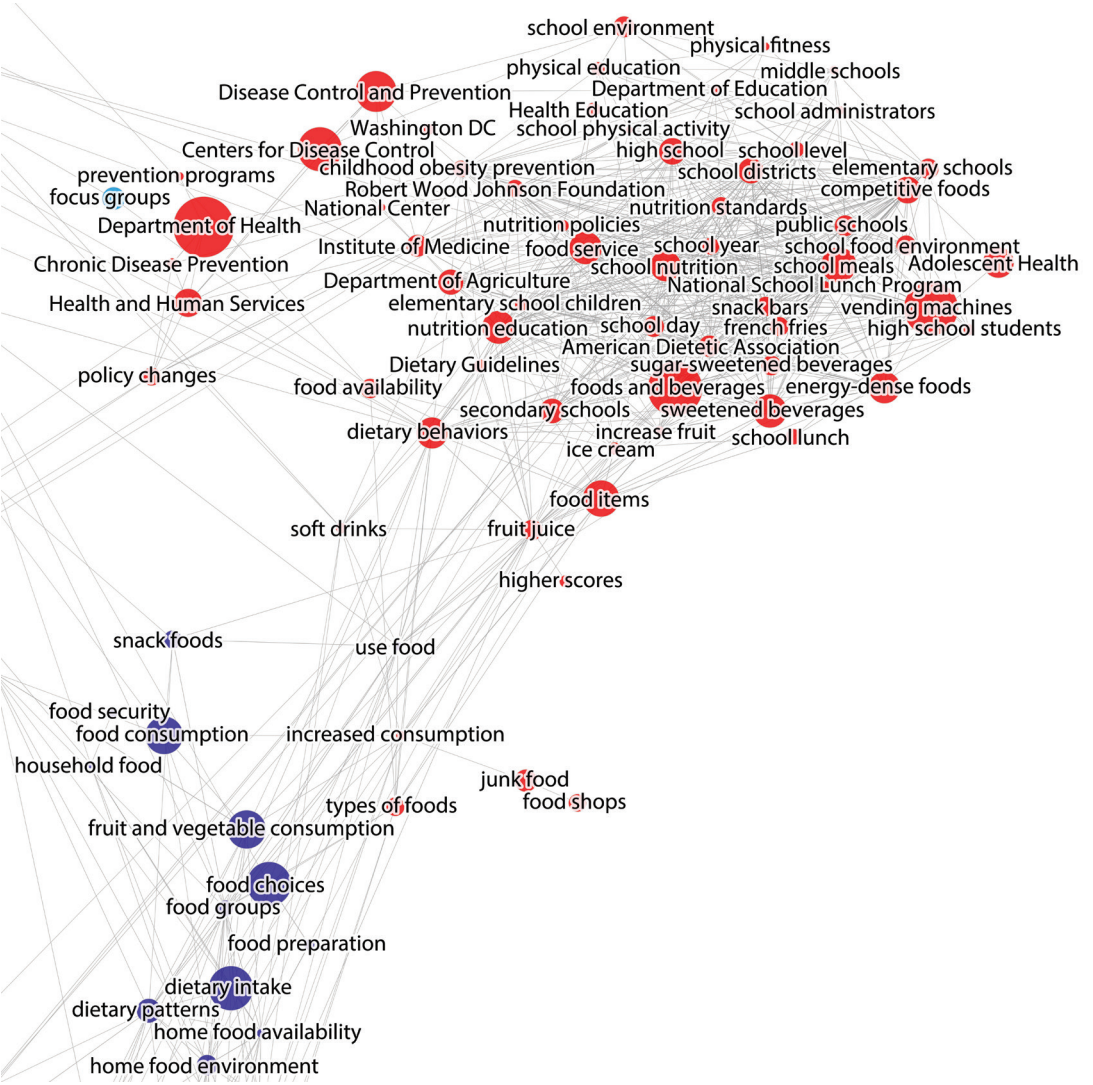
The second key practice in PDD, closely related to datafication, is the extensive use of flexible visualization. If a visitor dropped in on an ongoing PDD project, he would immediately encounter a number of data visualizations. On the walls he would most likely see pie charts, timelines, and large, poster-size prints of network maps showing hundreds of words connected by lines of various lengths and thicknesses. The network maps would be littered with comments on post-it notes, or the project participants may have written directly on the maps, adding comments, drawing arrows, or circling portions of the map. The visitor might also witness people gathering around laptops or larger screens more convenient for group discussions. Most of the time the screens would show steady images similar to the printed maps, but occasionally the participants would start some kind of operation that would cause the networks on the screen to slowly start moving as if the whole network was being pulled apart while some parts of it were still hanging together. Seconds later, the participants might stop the movement, study the new configuration carefully, and perhaps add color, zoom in, or in some other way manipulate the display. Occasionally, the participant would switch to an entirely conventional mode of data display: a spreadsheet with rows and columns.

There are two main reasons for data visualization's enthusiastic use by PDD projects. First, that visualizations under certain circumstances may have the power to be very *persuasive* and force participants to relate their discussions to a shared visual object. Second, that data visualization may open new opportunities for *participatory design*. Let us discuss the two reasons in turn.

The persuasiveness of visualizations has been a significant empirical topic in STS. Anthropologists and historians of science have investigated how natural scientists have developed the craft of producing images and of putting large amounts of information into condensed formats (Coopmans et al. 2014; Lynch and Woolgar 1990). Latour (1990) argues that visualization not only sorts and organizes material in a way that allows scientists to get beyond a data-overwhelmed situation; it also creates the powerful effect of equipping actors with objects, which they can mobilize

as allies in discussions with others who disagree. Accordingly, Latour describes the development of the sciences, their increasing persuasiveness and sophistication, as part and parcel of the invention of imaging and visualization techniques. To visualize, in the sciences, is thus not merely to illustrate something, which might just as well be represented in another format. To visualize is to draw things together in a way that produces clarity, strength, and persuasiveness. A pie chart showing the distribution of a national budget or a curve showing the correlation between smoking and cancer tends to be far more persuasive than anecdotes about the state of an economy or health risks. The persuasiveness of visualizations within science is to some degree the power or the magic that PDD projects tap into when they produce data visualizations. We can all be skeptical or simply refuse to believe specific studies or scientists. But we also all live in a world where the entire *genre* of scientific facts is heavily associated with graphs, maps, pie charts, tables, and the like. So when PDD projects deploy data visualizations, we are tapping into a well-established scientific genre that is generally known for its persuasiveness. More importantly, this persuasiveness serves a specific function as a boundary object among the participants in the PDD process. When participants have collectively agreed on an operationalization (of, e.g., authenticity) they are likely to take an interest in and try to relate to the visual outcome of that operationalization.

But persuasiveness of visualizations is not a controlled and predictable effect. On certain occasions we have found that data visualizations are almost more persuasive than we might want. Anders Munk found, for instance, that when he taught a project management course for engineers, he got far more attention from the listeners if he displayed a moving data visualization on a screen behind him. On other occasions we have discovered that the collective trust in particular types of visualization that STS scholars have described in branches of natural science is not necessarily present in the group assembled in a PDD project. The network graph, which is our most frequently used mode of data visualization, is challenging for many participants. Most people have strong visual literacy with respect to well-known forms of data display, such as a line graph or a pie chart. But these skills cannot be transferred to a network graph, which does not depict a correlation or a distribution but rather an elaborate set of relations between a large number of entities. Counterintuitively, the location of a cluster (up or down, left or right) does not carry any specific meaning on a network graph of the type shown in figure 5.1,[6] and neither does the specific location of the entities within a cluster. What does carry meaning is the number of clusters, their density, and the extent of connection between them. A part of our role when arranging a PDD project is to explain new formats of visualization and to guard against overinterpretation. This might be challenging when participants are eager to impute their previous knowledge of the project's field onto the maps—for instance, by quickly and directly associating the entities they know from the field with specific clusters on the maps.



5.1 A network visualization of key phrases in obesity research literature (excerpt from a larger visualization in Elgaard Jensen et al. 2018).

The possibility of establishing persuasiveness—despite the unpredictability just mentioned—is one benefit of visualization. The other benefit is the opportunity to arrange a *participatory* process. Participation springs from the *flexibility* of contemporary data visualizations. In contrast to more static images, data visualizations can be generated quickly, displayed and enhanced in various ways, and reworked swiftly. In a PDD project, visualization can do some of the same work that cheap, pliable materials such as cardboard boxes and Lego brick models did in the Scandinavian tradition of participatory design. The key to bringing designers and workers together was to use flexible, low-fidelity materials that made it possible to play around with

different future work situations and to bring both workers' and designers' knowledge into an open collective process. A crucial element of the collective learning in these projects was the quick feedback between exploring options, such as a particular arrangement of people and machines in a workflow, and the consequences of that arrangement, such as the possibility of assisting each other in the case of breakdowns in the workflow (Ehn and Kyng 1992).

What we try to achieve in PDD projects is similar. By rendering data in an accessible visual format, we put objects on the table, the wall, or the screen, which the stakeholders can immediately relate to their knowledge of the field: Does this make sense or not? Will this use of data be important and valuable to the field or not? Returning to our opening example, John and the other participants certainly asked themselves such questions when they were confronted with the discursive maps of Zika themes and decided to make filtering choices in their search for clusters of tweets that could be relevant to a citizen hearing on pandemics. Ultimately, the result of these choices was a series of new visualizations that gradually shaped the questions that the PDD project set out to answer.

We have come to encourage the quick development of alternative visualizations as a vehicle for creating learning situations that bring STS researchers together with knowledgeable actors from the fields we are exploring. Yet despite all the advances in computational speed and data handling, we have difficulties achieving the kind of participation that people might experience if they stand around a Lego brick model of their future office area and debate where to place the walls. In this scenario anyone can reach out and move a Lego brick wall and anyone can quickly respond to this future scenario. In a PDD project the flow of events is often more restrained: An issue expert might suggest a particular way to cut the data—for instance, by delimiting the view to a shorter time period or by filtering out the actors who are relatively unconnected to other actors. If the group decides that this would be an avenue to pursue, one of us would have to do some work to create the next data visualization. This may be a matter of minutes or hours, and it may require leaving the room to produce a new printout. For this reason we have found ourselves on the constant lookout for technical developments that enable even quicker iterations of data viewing.

MATERIALIZATION

The work with datafication and flexible visualizations makes it possible to build tangible outcomes whose equivalents did not exist before. Here are three brief examples: In a project on the topic of obesity research, the outcome was a prototype of an analysis showing a surprisingly fragmented discursive landscape in a large collection of journal articles (Elgaard Jensen et al. 2018). In a project with the Royal Danish Theatre, the outcome was a map showing how the active users of the theatre's Facebook page were also active on other culture-related Facebook pages (Munk, Madsen,

and Jacomy 2019). The map gave the theatre an image of what competed for its audience's attention. Finally, in a project on implementing a school reform in the municipality of Aalborg, one of the outcomes was an online resource (a database interface) that made it possible to search and visualize the connections between more than nine hundred specific visions for the future school (Madsen and Munk 2019). These visions were formulated at a large meeting by teachers, parents, administrators, politicians, and others, and the material remained available as a reference point for the ongoing discussions about school policy in the municipality.

As our three examples indicate, the outcomes of PDD projects may be quite different—depending on the participants' interests and the availability of data, time, and resources. Despite the differences, two general points can be made. The first is similar to the argument about flexible visualization: If the strength of PDD projects is that they bring different kinds of people together, then it is crucial to create situations in which their knowledge can meet and be negotiated. Looking at and modifying a data visualization together is one such occasion. The goal of producing material outcomes toward the end of a project is another way to propel participants to create something, which can be an object of negotiation and collaboration. Once a final visualization, a prototype of an analysis, or a user interface is on the table, the stakeholders can comment on whether this object would make sense to other actors in their field and, if not, how it might be improved. A material outcome here becomes a device for articulating the participating stakeholders' knowledge.

The goal of material outcomes raises a tension concerning the kind of democratic ideals that shape these outcomes. The tension can be described with reference to what John Law (2009) calls "constitutionalism" and "interference" (see also Munk and Abrahamsson 2012). Is our task to create a situation that is sufficiently messy to resist any attempt to subsume it under a singular ontology (interference)? Or is our task rather to account for the multiple positions from which we will eventually have to craft a common world (constitutionalism)? The answer varies from project to project. For instance, the prolonged collaboration with the municipality of Aalborg was premised on crafting a visualization of the public debate that was deemed legitimate across people with different opinions, whereas the sprint with the Royal Danish Theatre had a more experimental premise and was meant to disturb preconceived notions of the cultural audience.

Our second point about PDD outcomes relates to the reception beyond the direct participants. We have limited opportunities to trace the long-term fates of the prototypes we produce. But occasionally we do get the chance to witness how organizations use the results of PDD workshops. One such occasion arose when John—the person from our opening story—invited some of us to be present at the Danish Board of Technology's citizen hearing on pandemics. As always, the citizens were given an information leaflet but this time one that contained the materials (selected tweets) that had been identified at the PDD workshop. It was gratifying for us to see the

material printed in the leaflet. As far as we know, it was the first time that the board had solicited sources other than established experts. However, when sifting through the leaflet, we also realized that the usual amount of expert testimony was still in the leaflet, and it greatly outweighed the small handful of selected tweets. As the citizen hearing unfolded, we also realized that the board had organized the deliberative process with the invited citizens in exactly the same way as before and that the tweets were scarcely mentioned in the discussions.

In an almost ironic way, the tweets in the leaflet symbolize both the hope and the difficulties of PDD. It is no small thing to have facilitated the adoption of a new type of information source into the machinery of the Danish Board of Technology, and we don't think we could have done it if we had not involved John and other stakeholders in producing a material outcome with key intuitions of the board built into it. But it is also clear that anything we produce in a PDD workshop is at the mercy of many indirectly involved stakeholders and obdurate organizational practices. In the discussion that follows, we offer some further reflections on what this long run might entail.

DISCUSSION

So far we have discussed the enabling circumstances of PDD projects as well as some of the key practices in these projects. The practices suggest some of the visions that guide us, including our intention of making the most of the new opportunities for datafication, our ambition of working collaboratively with stakeholders, and our aim of establishing learning situations using flexible visualization of data. We believe it is fair to say that these visions have been fulfilled in some measure. We have indeed collaborated and learned together with stakeholders and we have made new things with large amounts of data. However, we have also encountered different effects—intended as well as unintended—that require further discussion.

The first is a surprising sense of empowerment as STS researchers, because our knowledge products seem to travel in new ways (see the introduction to this volume). Compared with the usual products of STS scholarship—articles and books—the digital products are often engaged with much more by different actors. This has launched us into the position of being the ones who can deliver visually appealing objects that attract a good deal of attention. Even more strikingly, we have experienced that collaborators are willing to grant power to digital maps—for instance, by describing the maps as the actual, data-based facts, as opposed to their own previous representations of the field, which the stakeholders begin to describe as being more loose, sense based, and incomplete. Finding ourselves in this situation as STS researchers is a highly ambivalent experience given our research tradition's long association with constructivism. It is clear, as Downey and Zuiderent-Jerak note in the introduction, that both "the scholar and the scholarship get made and done in the

process." Although we welcome the passionate engagement with the visualizations, we try to stress their constructed nature, the uncertainty of data, and the dangers of reading too much into them (Munk, Madsen, and Jacomy 2019).

A second effect, which almost runs counter to the first, is a sense of loss of control. Working with PDD projects means seeking out stakeholders and issue experts, which again means that STS researchers cannot occupy the role as the knowledge authority in the field in question. The stakeholders have a considerable say in framing the project and its outcome, and in the aftermath we have often seen that digital maps or devices have been carried away by the stakeholders and incorporated into arguments and decisions that would normally be beyond our reach. In a sense, our loss of control over the outcomes of the PDD projects is a success, because it fits the intention to engage stakeholders and produce something of use to their practices. However, it also means that we need to rethink not only our ideas of authorship and ownership of our knowledge products but also what sort of situated ethical responsibilities we have when these products are carried away by external partners.

Although appearing opposed, these two effects—the gain and the loss of control—also interact. When participants assign the quality of facts to PDD results, we as researchers are sometimes simultaneously disempowered and more likely to lose control, because the products travel much faster to other settings when wrapped up as facts. One example of this appeared in the school reform project with Aalborg municipality, which in addition to an interactive database of Facebook comments and posts, also produced a more static network visualization of the main themes and their association. The visualization suggested that not all teachers were united in their critique of the school reform, which politicians and municipal leaders could use to destabilize the opposition they were facing from the teachers' union. In this case, we as PDD researchers became very attractive collaborators for those already in a powerful position, something that empowers and betrays at the same time, because we appreciated the attention but did not necessarily agree with being mobilized as evidence against the teachers.

A third effect of PDD, which can be used to work against such dynamics, is the particular problematizing or even deconstructive effect that may be achieved through data work. We could use this as a slogan and say that the aim should be to *slow down reasoning by speeding up data*. Quick assumptions about, say, the unity of the field of obesity research or the typical views of parents on school reform may be challenged if one can collect and explore a large data set on the issue. But quick assumptions may also be challenged simply by *trying* to collect data, such as in the case of the Royal Danish Theatre participants who had to slow down and define authenticity before they could even begin looking for relevant data. The effect of slowing down reasoning[7] is in our opinion the feature that most strongly indicates how PDD is not a consultancy gig on the side but a mode of doing STS work (see the introduction). STS has a long tradition of working against received notions, of articulating multiplicity, and

of showing how things could be otherwise by attending to historical or ethnographic details. We think PDD projects—equipped with digital resources and in collaboration with stakeholders—instigate a similar kind of ontological inquiry that works against universalistic and too-quick assumptions about the world.

A fourth and final implication relates to our role as STS researchers, in particular our often self-appointed role as the ones who make multiplicity visible. When engaging in PDD projects together with actors who are profoundly involved in a particular issue, it becomes clear that the benefit of multiplicity is always a situated and contested matter. The kind of multiplicity wished for by one set of actors may run counter to the interests of others. In one striking example, we saw how a municipality's wish for a wide and broad debate about school futures effectively sabotaged the teacher union's attempt to defend the interests of its members by speaking in unison. Faced with cases like this, it seems that PDD projects not only empower STS and engage us with new and interesting collaborators; PDD projects are occasions for reflexive STS learning (see the introduction) and for reexamining core values of STS, such as the field's long-lasting devotion to multiplicity. In particular, PDD projects suggest that as we enter contested fields, we constantly encounter stakeholders who have a strategic interest in producing particular kinds of multiplicity.

This does not mean that we will simply become merchants of multiplicity hired by particular stakeholders to help them perform a targeted deconstruction operation. Datafication processes, such as those described earlier, can reopen entrenched normative commitments, both for researchers and for other participants who are faced with the very practical challenge of operationalizing goals and ideals. One way to stay in such troublesome moments is to make the data sets produced in PDD projects as widely available as possible rather than just publishing results and circulating selected visualizations that support particular points of view. In the Aalborg school project, for instance, we insisted that all the formulated visions should remain publicly accessible online. A related option is to prioritize visualizations that are interactive in ways that keep the analysis open for reorientation—for instance, by allowing future users to choose between alternative criteria for inclusion and different metrics for ranking data. Circulating data sets and making interactive visualizations open-ended harnesses the potential for slowing down reasoning that the multivalence of digital data offers and extends the making and doing of PDD to settings other than the initial one. Indeed, part of the trick of datafication is to realize that other situations and settings have already shaped what can be constructed as data, and in this spirit, it makes sense to anticipate more such situations after a specific PDD project comes to an end. One of the key tasks for developing PDD further is, then, to find ways to facilitate the ongoing making and doing of PDD. Ultimately, our ambition with the PDD projects is not just wrapping up projects with neat material outcomes but setting in motion processes that collectively engage more actors with datafication as a means for ontological inquiry and building new things.

ACKNOWLEDGMENTS

Our notion of participatory data design has evolved through extended dialogue and collaboration with a large number of fellow academics, students, and external partners—far too many, unfortunately, to mention and thank individually. We are profoundly grateful for the energy, resistance, and creativity that they have all contributed.

NOTES

1. The term was first used in a Danish book chapter (see Elgaard Jensen et al. 2017).

2. The following account is based on Anders Koed Madsen, Anders Kristian Munk, and Andreas Birkbak's experiences as participants in a project with the Danish Board of Technology. For a more detailed account of the project, see Birkbak, Madsen, and Munk (2020).

3. For reviews of the participatory design tradition, see Asaro (2000), Kensing and Blomberg (1998), and Törpel et al. (2009).

4. The five projects we refer to here are the following: obesity research (Elgaard Jensen et al. 2018; Munk et al. 2016), ASSET (Birkbak, Madsen, and Munk 2020), Aalborg school reform (Madsen and Munk 2019), the Royal Danish Theatre (Munk, Madsen, and Jacomy 2019), and fake news (Bach et al. 2018; Bounegru et al. 2017). Our thinking about PDD has also developed in dialogue with colleagues doing similar kinds of work, in particular Morten Misfeldt and collaborators' work on datafication of learning goals in Danish primary schools (Misfeldt et al. 2018). Digital methods development and reflections have also evolved through projects in which external stakeholders were not firsthand participants (Birkbak, Bornakke, and Papazu 2017). The PDD approach builds on broader efforts to involve stakeholders in digital STS, in particular the so-called data-sprint approach (Munk, Meunier, and Venturini 2019; Venturini et al. 2014, 2015).

5. Some authors consider datafication to be an inherently negative tendency in contemporary society. We use "datafication" as a neutral word indicating simply the process of gathering and relating data to a specific problem. Whether datafication is beneficial or harmful is an empirical question. For approaches similar to ours, see Flyverbom and Madsen (2015) and Schäfer and Van Es (2017).

6. The network graph in figure 5.1 is spatialized with a force-vector layout (Jacomy et al. 2014).

7. See also Whatmore (2009) and the way she draws on the Belgian philosopher of science Isabelle Stengers.

REFERENCES

Asaro, Peter M. 2000. "Transforming Society by Transforming Technology: The Science and Politics of Participatory Design." *Accounting, Management and Information Technologies* 10(4): 257–290.

Bach, Daniel, Anders Grundtvig, Asbjørn Fleinert Mathiasen, Asger Gehrt Olesen, Andreas Birkbak, Anders Koed Madsen, Anders Kristian Munk, et al. 2018. "Whack-a-Mole: En undersøgelse af falske nyheder og deres økosystemer." *Politik* 21(1): 41–58.

Bastian, Mathieu, Sebastien Heymann, and Mathieu Jacomy. 2009. "Gephi: An Open Source Software for Exploring and Manipulating Networks." In *Third International AAAI Conference on Weblogs and Social Media*, 8, 361–362. Chicago: AAAI Press.

Birkbak, Andreas. 2013. "From Networked Publics to Issue Publics: Reconsidering the Public/Private Distinction in Web Science." In *Proceedings of the 5th Annual ACM Web Science Conference*, 24–32. New York: Association for Computing Machinery.

Birkbak, Andreas, Tobias Bornakke, and Irina Papazu. 2017. "The Twitter-Thing: Retooling the Parliament into Issue Publics." Exhibition presented at the Data Publics conference, Lancaster, UK, March 31–April 2.

Birkbak, Andreas, Anders Koed Madsen, and Anders Kristian Munk. 2020. "Digital Methods Contributions to Citizen Hearings: A Techno-Anthropological Approach to Twitter and Technology Assessment." In *Technology Assessment in Techno-Anthropological Perspective*, edited by Lars Botin and Tom Børsen. Aalborg: Aalborg University Press.

Birkbak, Andreas, and Anders Kristian Munk. 2017. *Digitale Metoder*. Copenhagen: Hans Reitzel.

Blok, Anders, and Morten Axel Pedersen. 2014. "Complementary Social Science? Quali-quantitative Experiments in a Big Data World." *Big Data & Society* 1(2): 1–6.

Borra, Eric, and Bernhardt Rieder. 2014. "Programmed Method: Developing a Toolset for Capturing and Analyzing Tweets." *Aslib Journal of Information Management* 66(3): 262–278.

Bounegru, Liliana, Jonathan Gray, Tommaso Venturini, and Michele Mauri. 2017. *A Field Guide to Fake News*. Amsterdam: Public Data Lab.

Burgess, Jean, and Ariadna Matamoros-Fernández. 2016. "Mapping Sociocultural Controversies across Digital Media Platforms: One Week of #gamergate on Twitter, YouTube, and Tumblr." *Communication Research and Practice* 2(1): 79–96.

Callon, Michel, Pierre Lascoumes, and Yannick Barthe. 2009. *Acting in an Uncertain World*. Cambridge, MA: MIT Press.

Coopmans, Catelijne, Janet Vertesi, Michael E. Lynch, and Steve Woolgar, eds. 2014. *Representation in Scientific Practice Revisited*. Cambridge, MA: MIT Press.

Ehn, Pelle, and Morten Kyng. 1992. "Cardboard Computers: Mocking-It-Up or Hands-on the Future." In *Design at Work*, edited by Joan Greenbaum and Morten Kyng, 169–196. Hillsdale, NJ: Lawrence Erlbaum.

Elgaard Jensen, Torben, Anne Katrine Kleberg Hansen, Stanley Ulijaszek, Anders Kristian Munk, Anders Koed Madsen, Line Hillersdal, and Astrid Pernille Jespersen. 2018. "Identifying Notions of Environment in Obesity Research Using a Mixed-Methods Approach." *Obesity Reviews* 20(4): 621–630.

Elgaard Jensen, Torben, Anders Kristian Munk, Andreas Tamborg, Anders Koed Madsen, and Morten Misfeldt. 2017. "Participatorisk Data Design: En ressource for Capacity Building." In *Samskabelse og capacity building i den offentlige sektor*, edited by Hanne Katrine Krogstrup. Copenhagen: Hans Reitzel.

Flyverbom, Mikkel, and Anders Koed Madsen. 2015. "Sorting Data Out: Unpacking Big Data Value Chains and Algorithmic Knowledge Production." In *Die Gesellschaft der Daten: Über die digitale Transformation der sozialen Ordnung*, edited by Florian Süssenguth, 123–144. Bielefeld, Germany: Transkript Verlag.

Jacomy, Mathieu, Tomasso Venturini, Sebastian Heymann, and Mathieu Bastian. 2014. "ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software." *PloS One* 9(6). https://doi.org/10.1371/journal.pone.0098679.

Kensing, Finn, and Jeannette Blomberg. 1998. "Participatory Design: Issues and Concerns." *Computer Supported Collaborative Work* 7(3–4): 167–185.

Latour, Bruno. 1990. "Drawing Things Together." In *Representation in Scientific Practice*, edited by Michael Lynch and Steve Woolgar. Cambridge, MA: MIT Press.

Latour, Bruno. 2004. "Whose Cosmos, Which Cosmopolitics? Comments on the Peace Terms of Ulrich Beck." *Common Knowledge* 10(3): 450–462.

Law, John. 2009. "The Greer-Bush Test: On Politics in STS." Version of December 23. www.heterogeneities.net/publications/Law2009TheGreer-BushTest.pdf.

Lomborg, Stine, and Anja Bechmann. 2014. "Using APIs for Data Collection on Social Media." *The Information Society* 30(4): 256–265.

Lynch, Michael, and Steve Woolgar. 1990. *Representation in Scientific Practice*. Cambridge, MA: MIT Press.

Madsen, Anders Koed. 2013. "Web-Visions: Repurposing Digital Traces to Organize Social Attention." PhD diss., Copenhagen Business School Press.

Madsen, Anders Koed. 2015. "Between Technical Features and Analytic Capabilities: Charting a Relational Affordance Space for Digital Social Analytics." *Big Data & Society* 2(1): 1–15.

Madsen, Anders Koed, and Anders Kristian Munk. 2019. "Experiments with a Data-Public: Moving Digital Methods into Critical Proximity with Political Practice." *Big Data and Society* 6(1): 1–14.

Marres, Noortje. 2005. "Issues Spark a Public into Being: A Key but Often Forgotten Point of the Lippmann-Dewey Debate." In *Making Things Public: Atmospheres of Democracy*, edited by Bruno Latour and Peter Weibel, 208–217. Cambridge, MA: MIT Press.

Marres, Noortje. 2015. "Why Map Issues? On Controversy Analysis as a Digital Method." *Science, Technology, & Human Values* 40(5): 655–686.

Mayer-Schönberger, Viktor, and Kenneth Cukier. 2013. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. London: John Murray.

Misfeldt, Morten, Andreas Lindenskov Tamborg, Ane Qvortrup, Camilla Kølsen Petersen, Lærke Ørsted Svensson, Benjamin Brink Allsopp, and Lone Dirckinck-Holmfeld. 2018. "Implementering af læringsplatforme—Brug, værdier og samarbejde." *Tidsskriftet Læring og Medier (LOM)* 11(28): 1–21.

Munk, Anders Kristian. 2013. "Techno-anthropology and the Digital Natives." In *What Is Techno-anthropology?*, edited by Tom Børsen and Lars Botin, 287–310. Aalborg: Aalborg Universitetsforlag.

Munk, Anders Kristian. 2019. "Four Styles of Quali-quantitative Analysis: Making Sense of the New Nordic Food Movement on the Web." *Nordicom Review* 40(1): 159–176.

Munk, Anders Kristian, Mette Simonsen Abildgaard, Andreas Birkbak, and Morten Krogh Petersen. 2016. "(Re-)Appropriating Instagram for Social Research: Three Methods for Studying Obesogenic Environments." In *Proceedings of the 7th 2016 International Conference on Social Media & Society*, article no. 19, 1–10. New York: Association for Computing Machinery.

Munk, Anders Kristian, and Sebastian Abrahamsson. 2012. "Empiricist Interventions: Strategy and Tactics on the Ontopolitical Battlefield." *Science & Technology Studies* 25(1): 52–70.

Munk, Anders Kristian, and Torben Elgaard Jensen. 2014. "Revisiting the Histories of Mapping." *Ethnologia Europaea* 44(2): 31–47.

Munk, Anders Kristian, Anders Koed Madsen, and Mathieu Jacomy. 2019. "Thinking through the Data Body." In *Designs for Experimentation and Inquiry: Approaching Learning and Knowing in Digital Transformation*, edited by Åsa Mäkitalo, Todd E. Nicewonger, and Mark Elam, 110–128. London: Routledge.

Munk, Anders Kristian, Axel Meunier, and Tommaso Venturini. 2019. "Data Sprints: A Collaborative Format in Digital Controversy Mapping." In *DigitalSTS: A Handbook and Fieldguide*, edited by Janet Vertesi and David Ribes. Princeton, NJ: Princeton University Press.

O'Reilly, Tim. 2007. "What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software." *Communications & Strategies* 1. https://ssrn.com/abstract=1008839.

Peirce, Charles Sanders. 1955. "How to Make Our Ideas Clear (Selected Excerpts)." In *The Age of Analysis: 20th Century Philosophers*, edited by Morton Gabriel White, 143–153. New York: New American Library.

Rogers, Richard. 2009. *The End of the Virtual: Digital Methods*. Amsterdam: Amsterdam University Press.

Rogers, Richard. 2013. *Digital Methods*. Cambridge, MA: MIT Press.

Schäfer, Karin, and Mirko Tobias Van Es. 2017. *The Datafied Society: Studying Culture through Data*. Amsterdam: Amsterdam University Press.

Törpel, Bettina, Alex Voss, Marks Hartswood, and Rob Procter. 2009. "Participatory Design: Issues and Approaches in Dynamic Constellations of Use, Design, and Research." In *Configuring User-Designer Relations*, edited by Alex Voss, Mark Hartswood, Rob Procter, Mark Rouncefield, Roger Slack, and Monika Büscher. London: Springer.

Venturini, Tommaso. 2010. "Diving in Magma: How to Explore Controversies with Actor-Network Theory." *Public Understanding of Science* 19(3): 258–273.

Venturini, Tomasso, Axel Meunier, Anders Munk, Erik Borra, Bernhard Rieder, Michele Mauri, Matteo Azzi, et al. 2015. "Climaps by EMAPS in 2 Pages (A Summary for Policy Makers and Busy People)." https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2532946.

Venturini, Tommaso, Donato Ricci, Michele Mauri, Lucy Kimbell, and Axel Meunier. 2015. "Designing Controversies and Their Publics." *Design Issues* 31(3): 74–87.

Vertesi, Janet, and David Ribes, eds. 2019. *DigitalSTS: A Field Guide for Science and Technology Studies*. Princeton, NJ: Princeton University Press.

Whatmore, Sarah J. 2009. "Mapping Knowledge Controversies: Science, Democracy and the Redistribution of Expertise." *Progress in Human Geography* 33(5): 587–598.