December 4, 2020

A Mixed-Methods Study of Internet Standard-Setting

Nick Doty
UC Berkeley, School of Information

Status of This Document   *This is a chapter of a published dissertation:* Enacting Privacy in Internet Standards.

4   A Mixed-Methods Study of Internet Standard-Setting

This chapter describes the complexity of where, how and by whom Internet standard-setting happens and how I seek to study it, as an active participant, using semi-structured qualitative interviews and analysis of communications data. My goal is for this explanation to be useful both in interpreting the results of this research and in contributing to the research field of Internet governance.

For both participation and research, Internet standard-setting can at once be both surprisingly open and frustratingly opaque.

- participation is encouraged by: open process, extensive archives, ethos of individual participation
- participation is inhibited by: expertise, costs (time, money), history with a group

Setting the scene provides context for this openness and opacity with an example of the standard-setting work mode and its networked nature. The openness of participation and nonetheless the substantial barriers to it make for significant questions of the legitimacy of technical standard-setting as a model of governance or

regulation.[1] For the researcher, access to the standard-setting community reflects the tension of those participation trends. We have a rich corpus of participants, conversation and design, done in a relatively open and well-documented way. But there is also a maze of bureaucratic detail, technical jargon, pre-established personal relationships and outside-the-room activity, spread across many groups and hundreds of organizations in different sectors and geographic areas.

Studying standard-setting at different scales considers those properties of the technical consensus standard-setting process and the combination of methods for understanding the community and the implications for these multistakeholder processes in addressing debates over values such as privacy. Each of those scales is detailed in the following sections. To address that challenge, I have used my personal involvement and participation to inform my inquiries, interviewed a sample of participants privately to gain an understanding of their diverse perspectives and used automated analysis of mailing list archives to identify and measure broader trends across a larger community. Finally, I present the ethical framing of studying up and the protections for research subjects.

4.1   Setting the scene

I recall the first standards meeting I attended, showing up in no official capacity and with no particular affiliation.[2] Ten or fifteen people sit around a U-shaped table in a nondescript hotel conference room in Santa Clara, California; they are mostly white men, engineers at different tech companies. Someone at the front of the room is the chair, and projects a list of items onto a screen that the rest of the group faces. Other chairs are available, not at the table, around the edges of the room, where people less committed to this particular meeting can sit in, watch, maybe participate. I take one of these seats; at one of the breaks maybe I introduce myself to a couple of people. Everyone in the room has their laptop open. While there is a discussion happening in person, there is also an online chatroom – known to all the participants, to some remote attendees who couldn't make it in person, to some people who have an interest in this group but are currently in other conference rooms in the same hotel – with active discussion. At times, there is relative quiet in the room while everyone is typing and reading what others are typing. (Yes, this experience feels bizarre at first, but you get used to it.) Whenever

---

[1]See the earlier chapter Internet Standard-Setting and Multistakeholderism and Doty and Mulligan (2013).

[2]A vignette of a typical scene is provided to help the reader grasp the basic structure of this process; this is not based on field notes.

someone is talking in the room, someone else (the currently-designated "scribe") is taking notes on their statement, attributed to the speaker, in the IRC room.[3]

The conversation is loosely organized around a list of issues (that projection onto the screen in the room) related to the document being discussed:[4] the chair or editor asks a question about how something should be phrased or explained, and people in the room provide brief opinions; some back-and-forth, some through an organized person-by-person queue. The meeting is small and fairly casual; as the youngest, least experienced person in the room, I still feel comfortable chiming in on occasional points. Some of the questions are answered right here – after a brief back-and-forth, it seems like everyone is in agreement, the chair points that out and summarizes, the issue is "closed" and the resolution is recorded in the chatroom by the scribe. Someone else in the chatroom adds some notes with more detail, or types out who is responsible for implementing the change. But in many other cases, the question can't be fully resolved right here and now: someone needs to investigate a technical detail further, or an argument between two participants needs more fleshing out, and so it's noted that the conversation will be taken "to the list."

"the list" is the group's mailing list; a group like this relies on this piece of automated email-based infrastructure. This is by formal organizational policy in most cases: IETF and W3C create hosted mailing lists whenever they charter a new Working Group (or indeed many of the less formal groups as well). Meetings cannot be easily organized without such a broad, accessible communication channel, and a group without out-of-band asynchronous electronic communication can't do the discussion that makes up the work of a group made up of people living in different countries, working at different companies and participating either intensively or just occasionally. The mailing list's address is available on a corresponding web page about the group and widely advertised for feedback on any standards documents. Typically, the mailing list is public: anyone can subscribe, anyone who can convince the system they aren't a spammer can send a new message that will be distributed to the full group, anyone (subscriber or not) can read through a Web-hosted archive of every message ever sent to the list in the past. Messages range from very short ("+1") to thousands of words long. There are, to anyone not familiar with this kind of work, a lot of them. Long threads

---

[3]The minutes of this particular meeting are publicly available, recording most every statement, including my own questions.

[4]In this particular case, it was the draft specification of the W3C Geolocation API, being developed from 2008-2009 or so, that will enable web sites to ask to determine the user's precise latitude and longitude. Many W3C and IETF meetings follow a similar agenda pattern.

with very detailed arguments about any issue considered by the group, automated or bureaucratic messages describing technical changes or distributing the minutes of past meetings, casual personal conversation or complaints about some piece of software or another that are met with a reminder about the intended scope of the mailing list. List conversation can be friendly and informal or, at least as often, brusque and insulting. Prior to attending that conference room "face-to-face," I'd sent a few messages to the group's list; some would receive responses from an interested party or someone trying to gather support behind a particular idea, others would be ignored.

But even looking at these two "sites," distributed and wide-ranging as they are, would be too blindered to understand all the points of conversation between the formal participants of a standard-setting process, much less to capture: the debates within organizations; the business and policy discussions between firms inside and outside the same field; the relationship of companies and regulators in different jurisdictions; or the effects of software out in the world and how it's used. I recall describing this research project on privacy in standard-setting to an important policymaker (a non-participant stakeholder, in the terms of this research) who asked me, pointedly, was I going to limit my investigation to the standard-setting groups themselves? (Someone else in the room quietly shook their head "no," urging me to avoid my obvious blunder.) Trying to play the good academic, I responded with something about focusing research on a limited scope for the purposes of finishing a dissertation; this was quietly received as a sign of my apparent obliviousness.

In laying out the illustrative cases in this work, I have followed discussions that formally take place in standard-setting fora, but also describe interviews with participants and non-participant stakeholders and cite relevant press writing and other announcements broader than just the working groups.

4.1.1 The networked site   While not traditional in the sense of historical anthropology, this is also not an entirely novel environment. For example, Coleman has written about the free and open source software movement by researching at week-long conferences and reading extensive mailing list archives (Coleman 2012); Some Internet ethnography has tried to focus on the properties of the virtual site as a place (relying on metaphors of "cyberspace," for example), the communities that exist in that place and how to conduct research "in the field" in those settings (see lists of citations in Davies 2012). But more relevant to the distributed, wide-ranging and on-line/off-line communication styles of standard-setting is

Burrell's argument of a networked site with multiple entry points and connections (2009), Kelty's view of "distributed phenomena" (2008) and Hine's ethnography of mediated interactions and the "richness and complexity of the Internet" (2000).

Rather than marking a bright and arbitrary border around the site, I have tried to follow the participation in the distributed Internet standard-setting process *where* it happens – ranging from formal in-person meetings of groups with a specific membership, as well as teleconferences and mailing lists – as well as *who* is involved – from formal leadership to those non-participant stakeholders who observe or influence without being in the room. At the same time, to focus my inquiry, I have tried to focus on some core settings and then expand out to more peripheral involvement to supplement that study. Following the cases described in Privacy and Security: Values for the Internet above, I have set as a core group, the Tracking Protection Working Group who debated Do Not Track standards at W3C. While that group does have, in some ways, a formal membership list, meaningful participation also expands out to people who joined teleconferences, in-person meetings or mailing list conversations, those non-participant stakeholders who followed the process or influenced it in some ways, and further out the casual observers or even just the affected parties who had no awareness of it. By necessity, different stakeholders and sectoral groups will be larger outside the group than they are internally, even where represented internally. In the DNT work, some groups have deeper core participation (for example, advertising companies and consumer advocates) while others (for example, policymakers) are more peripheral. These wedges of depth of participation extend outwards: another way to picture participation.

The debate over encrypting Web traffic doesn't have as singular a core standards body locale. The same debate was ongoing in specific working groups, security area groups, in IETF plenaries and within and between companies involved in implementations and deployments of software at different layers of the stack. Even so, we can see conversations happening in Internet and Web standard-setting at the center of a set of concentric circles that encompass firms, advocates, policymakers and users. There is overlap in the participants in standardization around Do Not Track and standardization around Web encryption even though the technical details of those projects are quite distinct.

This diagramming is not intended to value the importance of one group or one setting over another; as biased and interested in the impact of standards as I am, I still wouldn't call it the most important part of the development of the Internet and the Web. Indeed, research subjects are explicit in denying the framing of standard-setting as the most essential step and public writing from standards
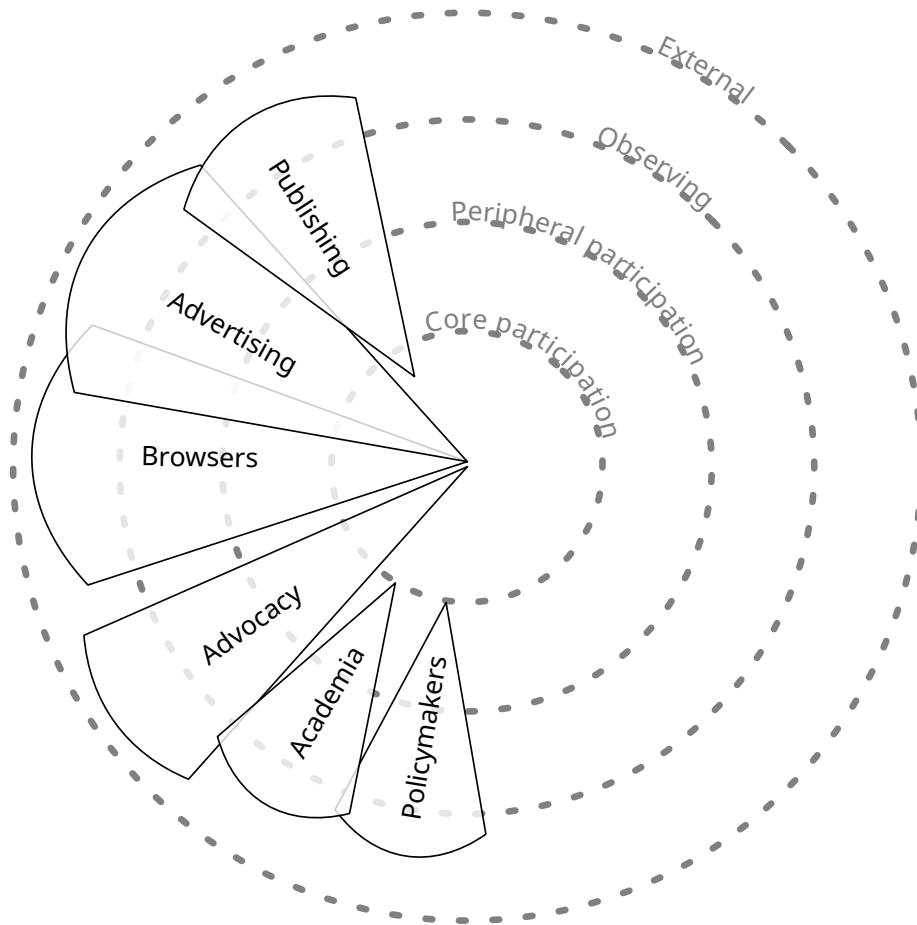
Figure 1: Sketch of core to periphery of stakeholder groups around Do Not Track, showing "wedges" of participation.

experts also emphasizes that point, as described in Internet Standard-Setting and Multistakeholder Governance. Instead, these diagrammatic maps of the standard-setting process show how this study of technical standard-setting processes is situated and how these distributed working groups are connected to others.

4.2   Studying standard-setting at different scales

In studying Internet technical standard-setting, even scoped to Internet and Web consensus standard-setting around the values of privacy and security, I am faced with the challenge of grappling with this diverse, distributed, networked site.

In order both to validate findings from intensive qualitative investigations and to identify the character and causes of apparent trends present in quantitative data, this study takes a mixed methods approach to investigation. This site involves, significantly, individuals with their lived experiences, dynamic interactions between people, and organizational structures that connect large industries. To encompass those scales, this research also includes different methods to gather insight at different scales of study. Different research projects can contribute by including just one of these methods, but my central argument focuses on the interaction between individual participation and larger organizational settings and so this work attempts to examine those different components.

Table 1: Applying different methods to different subjects and scales of study

Methods	Subject of study	Scale
self-reflexive	lived experiences	micro
qualitative	interpersonal dynamics	micro/meso
quantitative	organizational, high-level trends	meso/macro

How can these different methods at different scales inform one another? Here are three examples from my ongoing research:

identifying areas for closer investigation Quantitative analysis of a large dataset like the full corpus of IETF mailing list archives allows measurement of social network properties like closeness centrality to identify important figures that are highly connected to different subgroups.[5] Identifying central people

---

[5]See this research notebook on IETF participation for the methods, data and initial conclu-

or groups in a large community can help identify key people to interview or establish what organizational roles are worthy of more intensive observation and inquiry.

validating and measuring disparities My personal experience at standards meetings prompted the question of demographic imbalances in standard-setting group participation, observing male-dominated in-person discussions or apparent reticence of participants from East Asian countries. Quantitative estimates of gender in participation on mailing lists provides a point of comparison (do computer-mediated communications have the same effects?), a way to validate (does the disparity apply across multiple contexts?) and to identify potential variables that might affect the distribution (do some work areas show less demographic imbalance?).

explaining the effects of interventions Explaining interventions is an especially challenging task for research. Quantitative analysis can provide comparability, but suffers from confounding factors or wide variations in interpretation. Qualitative research can provide rich description, but shies away from causal explanations or broad external validity. By using both quantitative analysis of the trends across hundreds of documents and a qualitative understanding of reading documents and talking to authors we can better explain and contextualize the effect of mandates and guidance on the presence and significance of security considerations sections in standards (Doty 2015).[6]

Subsequent sections of this chapter describe the specific methods and tools used to gather and analyze data from different sources, roughly grouped in the same order of self-reflexive, qualitative and quantitative methods described here. Integrating those methods as I suggest[7] is an ongoing challenge for me, but one where I hope to make a contribution.

---

sions from calculating closeness centrality: `https://github.com/npdoty/bigbang/blob/ietf-partic ipation/ietf-participation/IETF%20Participants.ipynb`

[6] While this has been a part of my research practice on standard-setting, as in the cited paper, this dissertation work does not report on measurements of the effects of interventions.

[7] This model of mutually-informing scales of study was also presented at the Protocol to the People event at the Turing Institute (Doty 2018).

4.3 Researcher position

As a participant as well as a researcher, I aim to use my own perspective, including the challenges of working in a diverse multistakeholder setting. I have approached privacy in standard-setting as an active participant myself — not a putative detached observer — and that will inevitably be apparent in my work. Research requires reflexivity (Watt 2007) about the research process, my own subjective experience and the effects of my involvement. While "reflexivity" is used in many different ways, what I intend here is methodological reflexivity (Lynch 2000) — awareness of my own beliefs and experiences (because they are unmediated) and awareness of my multiple roles within the groups of study themselves.

Not unlike many subjects of this research, I have held many different roles (at different times and simultaneously) with respect to development of the Internet and World Wide Web. I have been a user of the Internet and the Web since I was first given a one-hour tutorial at my local public library in small-town Virginia (circa 1992) and managed to discover a Star Trek fan page. Like many in my generation, my technical training was mostly self-taught; in middle school (the mid-90s) I started a web design "business" with a classmate (we never had paying clients, although he later founded an online community called Reddit), and learned enough HTML and JavaScript to make images change on mouseover. While computer science classes were accessible in high school and college, they never covered Web technology; I taught myself PHP and explored the privacy violations present by running `tcpdump` on misconfigured switched networks. After a short stint in software engineering at Microsoft, I first encountered Web standards as a Berkeley graduate student, following mailing lists and writing workshop position papers and research reports regarding geolocation privacy. From 2011 through 2015, I was employed as part-time staff at the World Wide Web Consortium (with "Privacy," rather than a job title, on my business cards), managing the Tracking Protection Working Group (TPWG) and the Privacy Interest Group (PING), work that ranged from being a job for one to two days each week up to entirely consuming all my waking thoughts. As staff, I recruited participants, handled process and logistics, organized meetings and events, provided technical support and responded to press inquiries, but also participated in the discussion, debate and design work of the groups themselves. In 2013, I started editing a guidance document for mitigating browser fingerprinting (2019) and in 2014, I took on the role of Editor for Tracking Compliance and Scope in DNT (2019), roles that I continued as an unpaid volunteer after 2015.

Throughout these different roles, I also identified myself as an academic, a

researcher affiliated with UC Berkeley with an interest in privacy and technical standards, and continued (where I could find the time) to publish research in journals and workshops. Having those multiple, overlapping roles (user, amateur Web developer, W3C Staff, academic, Team Contact, Editor) complicates my experiences and my position. But such role diversity is also not exceptional among standard-setting participants themselves, who switch employers, job titles, rhetorical positions and stakeholder groups while maintaining a connection to a standard-setting body.

Having a stake in the outcome might seem like a violation of the neutrality expected from:

a) a researcher,
b) someone writing a standard for compliance, or
c) the staff of a standard-setting body.

In each case, I believe that purported detached neutrality is neither plausible nor constructive to the aims of the commons.

A common point of confusion in explaining the standard-setting process to press was the apparent contradiction between standards both regulating the behavior of, and being debated and developed by, the implementers of those standards. "Isn't it like the fox guarding the henhouse?" This question arose most often when noting that initial editors of DNT specifications included representatives of Google and Adobe, but the question and confusion applies to the process of developing specifications more generally, especially because editors do not typically have ultimate decision-making authority. Consensus standards can function not despite but *only because* they are developed by the impacted groups that are the implementers.[8] For practices to be voluntarily adopted and practically informed, it is constructive to have stakeholders as authors and collaborators. Scholars have argued that standard-setting is possible only because of the shared motivation towards a common goal (Cargill 1989).

Standard-setting bodies vary in the roles of staff and the different kinds of neutrality that they practice. W3C itself has a mission, the apt, if anodyne: "to lead the World Wide Web to its full potential by developing protocols and guidelines that ensure the long-term growth of the Web." Staff are expected to share that goal

---

[8] See "The Consensus Standard-Setting Model" in Internet Standard-Setting and Multistakeholder Governance. That some impacted groups are necessarily well-represented, however, does not mean that all impacted groups are, which is a key question for the legitimacy of techno-policy standards (Doty and Mulligan 2013).

and work towards that mission, which is explicitly non-neutral. But there is an expectation that the Consortium is neutral with regard to the members, many of which are direct market competitors. Staff have very different backgrounds and participate to different degrees in the groups they facilitate, but have considerable discretion within the broad mission and can take vocal positions in standardization discussions. This is to some extent inevitable and to some extent a particularity or historical tradition of Internet standard-setting: that individuals have perspectives and express them in a way that is distinct from organizational priorities.[9]

Finally, as a researcher, deep involvement with a particular perspective seems contrary to the traditional position of a detached observer as in the model of the pith-helmet-wearing anthropologist documenting primitive tribes.[10] There is ongoing qualitative research on technical standard-setting communities using participant observation and interviewing from researchers observing these groups; I look forward to seeing their results. However, detachment can sometimes assume an impartiality that does not exist, where reflexivity allows us to recognize our positions as researchers (our personal perspectives, our effect on discussions, how our methods respond to community dynamics). In terms of access, rapport and nuanced understanding, I believe there are distinct advantages that my deep personal involvement can bring. Mine is necessarily a unique perspective, and one I hope can bring the reader more deeply into a complex setting.

My own subjective experience can be an advantage in connecting with subjects in the community and understanding their lived experience (Sandelowski 1986, discussing the advantages of subjectivity for confirmability). My first-hand familiarity with the challenges of multistakeholder participation has been useful in establishing rapport during interviews and in identifying (via experience) and confirming (via interviews and conversations) emotional responses present in the community. However, because my own positions are clear to stakeholders, that could inhibit candor where a participant chooses not to expose disagreements. In addition, where stakeholders have identified me as antagonistic, that may limit access altogether. As a researcher working on privacy, my position that privacy is an important value to be supported in the design of the Internet and the Web has

---

[9]For more on the distinctive role of the individual in Internet standard-setting, see: "The Internet and Requests for Comment" in the chapter Internet Standard-Setting and Multistakeholder Governance, "A network of actors and actions" in the case Do Not Track, a "handoff" and Individuals vs. organizations in the findings.

[10]For some audiences in the area of qualitative research, the absence of a neutral detached position is so widely accepted as to need no explanation, while for others this assumption may be important to rule out.

been well-known, and while that view might be common in the abstract, there are surely other participants who identify me as too focused on the value of privacy, or not focused enough, compared to other values.

I stay aware of the limits of relying on my own experiences, lest the study become confessional.[11] Personal experience can seem all too vivid, but lacks the replicability or investigability of rigorous, systematic research. In studying these multistakeholder groups, which almost by definition include people with very different backgrounds and perspectives, using my own experiences as canonical would be a mistake. I have tried to use personal experience in the form of vignettes or self-reflection to illustrate a perspective and to illuminate settings that may be unfamiliar, but rely on qualitative and quantitative methods to analyze the community and processes. Noting the issue of reflexivity does not automatically deflect all concerns, but this awareness should exist throughout my presentation of this research and in evaluation of my methodology.

4.4   Interviewing

Even where documentation is extensive and participation can be directly observed, understanding the internal views of members of a community can be difficult. To gather insight into that emic perspective, I have conducted semi-structured interviews with various participants (and some non-participant stakeholders) in Internet standard-setting.

These interviews were guided to gain insight into both the people and their personal perspectives and the standard-setting process and how they perceived it.[12] I began interviews with questions about participants' backgrounds and their roles in their organization; I inquired into personal views on privacy – how they define it, what kinds of privacy concerns they identify; and then I asked them about their experience with technical standard-setting, when privacy came up in those conversations and how they saw their role. For those involved in the Do Not Track process in particular, I asked about their particular goals for that process, how they perceived debates that took place and how they viewed other participants.

---

[11]"Here is just one example of the total wrongness of something I tend to be automatically sure of: everything in my own immediate experience supports my deep belief that I am the absolute center of the universe; the realest, most vivid and important person in existence." — Wallace (2005)

[12]The full interview guide I used for privacy in the standards process is included in the appendix.

I conducted 27 interviews over the course of this project, not evenly spaced between late 2012 and late 2019. That distribution was largely driven by this researcher's varying time that could be dedicated to data collection as my direct participation and employment took up less space. As a result, the interviews don't attempt to show a comparative assessment at a single snapshot in time, but include both ongoing and retrospective viewpoints.

4.4.1 Dimensions for sampling Understanding the perspectives of a diverse community, or a community that is at the intersection of various groups rather than a single cohesive or homogenous setting, provides challenges for the sampling process. While we cannot *a priori* know all the variety among our potential research subjects, we can sample in a way informed by theoretical considerations. Among those: multistakeholder groups specifically prompt the question of how different stakeholder groups are represented and operate in such a process.

Can we neatly divide the participants of, say, the Tracking Protection Working Group into a clear faceted classification of distinct stakeholder groups? No, but we can nonetheless identify important apparent distinctions. Mapping different and overlapping stakeholder groups is feasible based on my working experiences with Web standardization and analysis of membership lists. We might also profitably use mailing list participation as a proxy to confirm or expand the representation of different groups, to the extent that we can evaluate affiliation. Research subjects can confirm or reject that sampling frame. While I have access and understanding of different participant groups to sample from, I also ask participants to recommend particular people to talk to as a form of "snowball sampling." This is done less for convenience and more to get the participants' own views of what groups or perspectives might be missing; in order to maintain the privacy of participants from other participants in the study, names are asked for without revealing who has already been interviewed.

For W3C standardization as a whole, I have tried to represent in this diagram the different overlapping groups of stakeholders, their intersections of what they represent and where they align and their levels of participation and influence in W3C processes. This working sketch is based on a review of W3C organizational membership and my personal experience with W3C organization and working groups; as such it is only one personal view out of potentially many different ones, but my hope is that it could give outsiders some idea of what sectors are present.

In addition to sampling different stakeholder groups, demographic differences in participants is important for our ethnographic study to explore the effects that
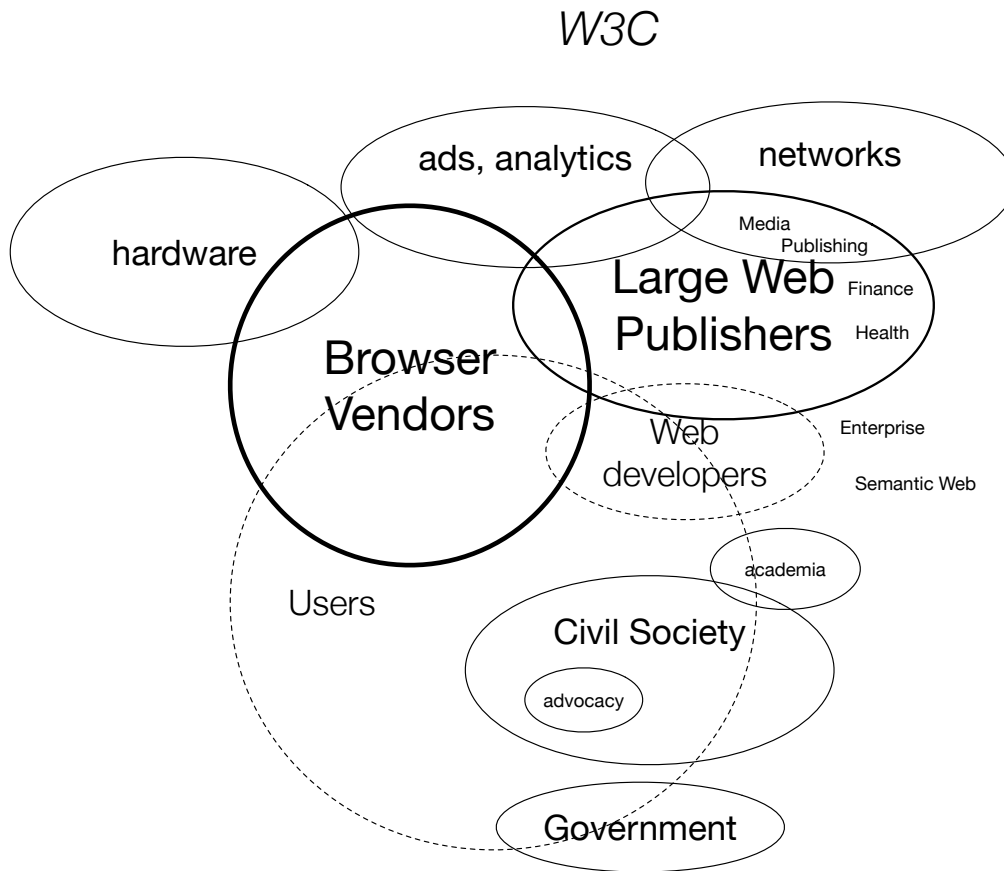
W3C

Figure 2: Mapping of overlapping stakeholder groups at W3C.

standard-setting process might have on demographic representation or the partic-ipation of different subgroups. There are many demographic dimensions that may be relevant to questions of legitimacy over the design of Internet protocols. Be-cause tech communities face prominent controversies over sexual harassment and discrimination in employment contexts, gender is one such area of interest. Only 15% of my interview subjects were women: similar to the gender balance in tech firms and in Internet standard-setting groups, but still very far from proportional.

Because this is a theoretically-informed sample, we might also choose to oversample certain groups of potential importance. As leadership was a theme identified early (from interviewees not in leadership positions), this research has tried to particularly include those with formal or informal leadership roles.

While studying a particular standard-setting working group may allow for a precisely-defined scope of membership, this study also seeks to include non-participant stakeholders and peripheral participants. Peripheral participants – sometimes involved and sometimes not, or attending meetings but not vocal – might provide an outside perspective and insight on why people choose not to participate.

Finally, sampling can be explicitly used to mitigate biases, either in the researcher or in potential access. One such danger is that as a non-confrontational person myself it might be especially easy to speak with a significantly skewed sample of people who are generally agreeable or who share perspectives or goals with me. Statistical representativeness is not a goal of this method, but not sampling at all from entire subgroups with a particular, distinct perspective would harm the breadth of understanding.

There is a subtle but important distinction between being more-or-less friendly to work with and being more-or-less supportive; indeed, these might be orthogonal dimensions. For my own purposes in identifying this variety in stakeholders and confirming the intentional diversity of my sampling, I developed this two-by-two matrix of Tracking Protection Working Group participants based on my own experience. For this purpose, I believe using my own experience to be especially apt, as I'm attempting to counter any internal preference for similar or agreeable subjects.

For the dimensional axes: the vertical axis shows the spread from being "friendly" (easy to work with, e.g.) to "difficult" (more likely to have unpleasant interactions); the horizontal axis is between "supportive" (shared the basic goals of the process) and "antagonistic" (opposed the existence of the process or the stated goals). There are certainly people in the friendly/supportive quadrant: people bought in to the process and happy to collaborate. And the direct opposite quadrant is also easy to identify: people who were opposed to the Do Not Track process in every way and who seemed to personally dislike me as well. This group may be difficult to access. That the dimensions are orthogonal (or at least substantially distinct) is supported by the presence of people in the other quadrants. A substantial group of people were regularly professional and nice, always happy to talk about professional and personal topics but were nonetheless substantially opposed to Do Not Track or to W3C processes. And while there may not be the most extreme examples, there were also people known to be difficult to work with who were nonetheless significantly supportive of, or potentially aligned with, the standard-setting process. As reported in the Findings, interviewees confirmed these dimensions as orthogonal and even posited similar additional ones about
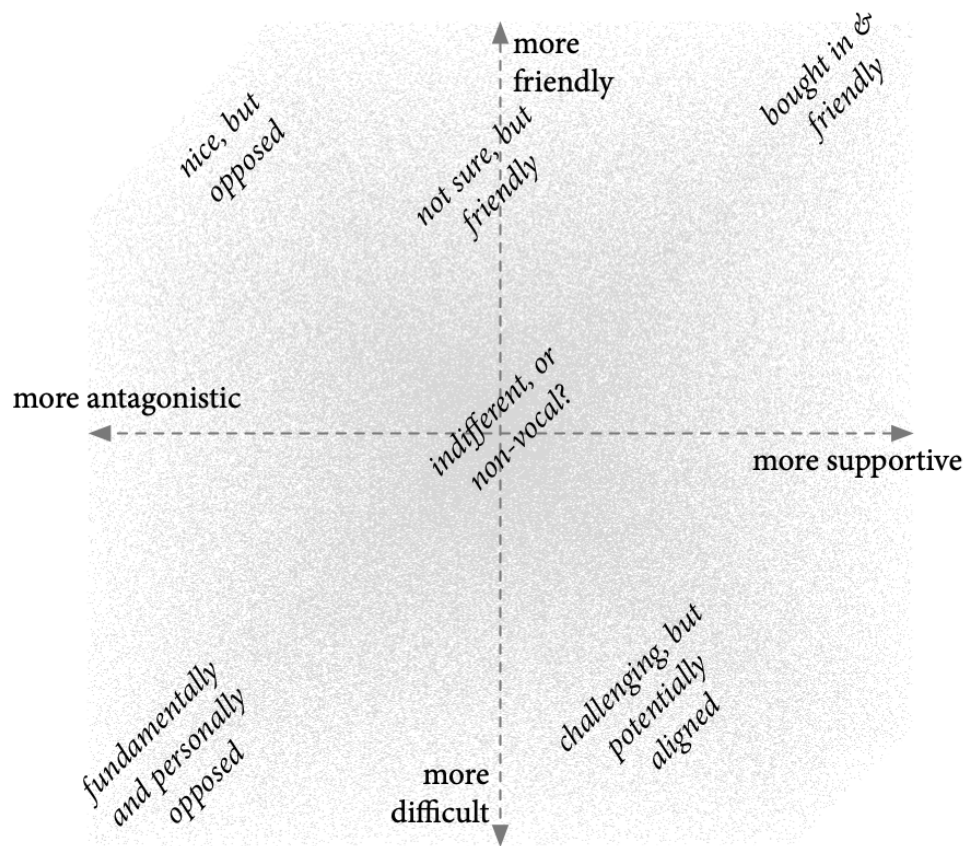
Figure 3: Sketch of distribution of orthogonal dimensions of difficulty-antagonism, with personal notes for the DNT standardization process.

Smaller sample for saturation

Capture themes
Homogenous population
Iterative sample
Thick data
Concrete codes
Stable codebook
Code identification

Larger sample for saturation

Develop theory
Heterogeneous population
Fixed sample
Thin data
Conceptual codes
Emerging codebook
Code meaning

Purpose
Population
Sampling Strategy
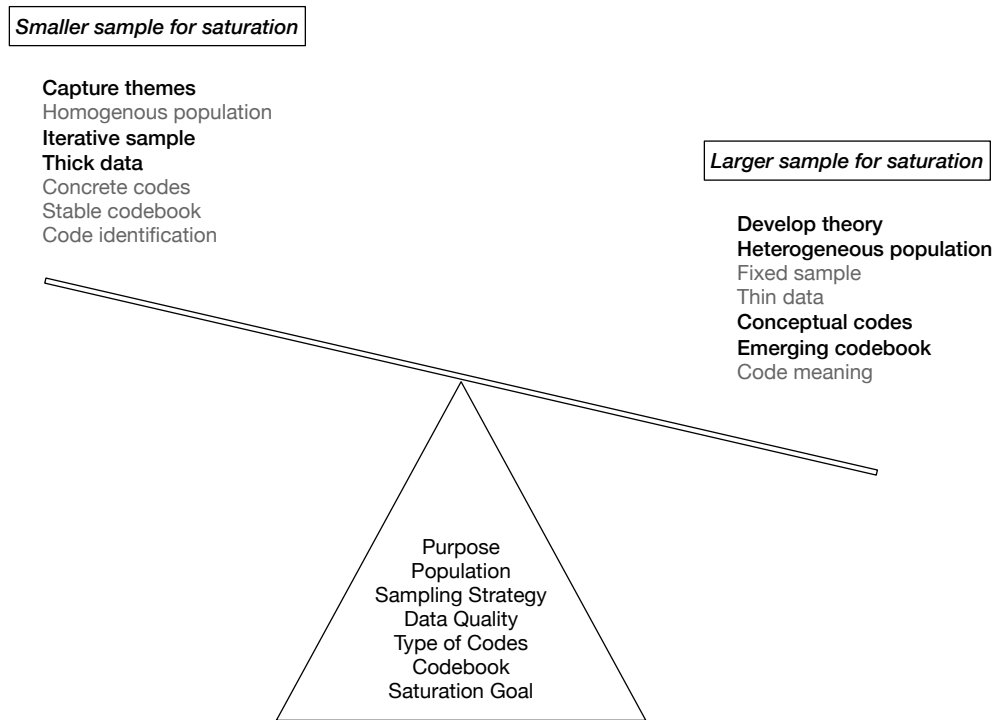Data Quality
Type of Codes
Codebook
Saturation Goal

Figure 4: A modification of the Hennink, Kaiser, and Marconi (2017) parameters of saturation diagram, highlighting the particular factors used in my study.

participating in good or bad faith.

Beyond those dimensions of diversity, I've attempted to collect data up to a point of meaning saturation (Hennink, Kaiser, and Marconi 2017).

Regarding the saturation parameters, this work has weights on both sides of the scale, as highlighted in the diagram. The heterogeneous population; emerging, conceptual codes; and theoretical interest all suggest a larger sample. But having access to thick data (and other sources of data, although this mixed and multi-scale method isn't considered) and iterative sampling allow a relatively smaller sample for saturation.

4.4.2 Coding, memoing and writing through Interviews with these participants, who have been largely engaged and candid with their experiences and expertise, provides an incredibly rich corpus, reflected in the transcripts that pile

up in encrypted disk images. So much is covered, and there are so many threads to pull at, that it can be overwhelming.

I've followed a practice of initial coding and focused coding (Strauss and Corbin 1990; as cited by Lofland et al. 2006) to identify common themes and important contrasts from that larger corpus.

In addition to coding to capture the topics and terminology raised by participants, I want to capture also the holistic perspectives that I hear from subjects. To do this, I write brief (approximately 1-page) memos after completing coding an interview, sometimes supplemented by the handwritten notes I took during the interview, to describe the ideas that stood out to me – a particular motivating argument or novel idea, say – as I reviewed the interview experience.

Without trying to replicate every point that the research subject is making, memos provide a brief but more open-ended mechanism to collect key points or insights that might not be easily represented as a single term or phrase for a code. At this slightly higher level, I can write down perspectives that might not be obvious in the interviewee's words but still come through in my reading of the conversation. Theoretical memos capture "momentary ideation" (Glaser 1978; as cited by Lofland et al. 2006) of my own thinking on reviewing an interview during the coding process.

To draw important insights from the corpus of interviews and their associated codes (hundreds after the multiple rounds of coding), I've identified common themes that appear in codes across multiple interviews and clusters of related codes into key themes to drive deeper analysis and my write-up of findings. In reviewing the quotes from different interviewees that touch on that same theme, I can get a sense of the variation and pull out quotes that are illustrative of a typical viewpoint (expressed several times by different people) or a distinctive viewpoint (one that sharply contrasts with other views).

As this methodology is qualitative, I don't have the statistical backing to show numerically that one view is most common or that some arguments are significantly more widely held than others. Instead, my goal is to show the existence of important themes among the participants and non-participant stakeholders and then to use the subjects' own words to help describe that experience for the reader.

4.5   Analyzing communications

Internet standard-setting is rich with communications artifacts: mailing lists, meeting notes, drafts, revisions and published documents. In this work I have focused on group mailing list archives for communications data that show group

discussion and interactions, but there is a real opportunity for our research field to investigate these traces as a supplementary method more generally.[13]

4.5.1   Mailing list analysis   More than any other single "place," Internet standard-setting has happened on mailing lists. While not contained in a single geographic location, the mailing list functions in a place-like way, along the lines of what we see in critical geography's analysis of place as opposed to space – as Massey argues, made up of social interrelations and flows of people and communication (1994).

There are other places of interest – important decisions are made in face-to-face meetings or persuasive conversations that happen in private settings at an office or bar – but the majority of argument, debate, discussion, positioning, presentation and reasoning in these groups takes place in email fora.[14] These lists are places in the sense of containing and mediating these interactions, even though the participants are geographically and temporally dispersed. That these mailing lists are typically publicly, permanently archived makes them rich sources of retrospective study and analysis.

4.5.2   BigBang

> BigBang is a toolkit for studying communications data from collaborative projects.[15]

Collaborators at UC Berkeley, Article 19 and the University of Amsterdam have developed BigBang as a collection of tools for analyzing traces from open source software and Internet governance groups; this collaboration is supported by DATACTIVE, a research group focused on data and politics. Independent researchers pursuing their own projects have collaborated on this tool because of commonalities in the communications tools used by these software development, standards development or decision-making groups – all typically use archived mailing lists to develop community and discuss their work. Identifying who participates, how they participate and how the structure of these groups affects their work is valuable to our group of social science researchers even though – or

---

[13]See, for example, the concept of trace ethnography (Geiger and Ribes 2011).

[14]This is a contingent, historical conclusion rather than a necessary or normative one. There are some groups that use chatrooms for a larger fraction of discussion and recent use of GitHub and software-development lifecycle tools for issue tracking are becoming more common in Web standards work.

[15]https://github.com/datactive/bigbang

perhaps especially because – different collaborative communities use these online communication tools in distinct ways.

BigBang has been used for collection and analysis of mailing list archives and Git version control repository information. Functionality has been developed for the following forms of analysis (with examples of specific measures or data considered):

- traffic analysis (messages over time)
- demographic analysis (gender, affiliation, country of origin, etc. of participants)
- social network analysis (centrality, connectedness, assortativity)
- content analysis (trends in word usage)

Those different forms of analysis allow for responsiveness to different classes of research questions. My hope is that we can learn from practice what kinds of data is appropriate to what kinds of research questions and what the practical challenges and feasible solutions are in studying mailing list data. That kind of work can set researchers up for a wider range of future projects along these lines.

Traffic analysis can illustrate patterns of activity: that might include trends across standard-setting fora altogether or the typical pattern of a working group. That activity can also show community responses to exogenous events, as I've explored with privacy and security activity after especially relevant Snowden disclosures about Internet surveillance (Doty 2015).

Demographic analysis can provide evidence on who is participating, relative levels of participation between different subgroups and how demographic variation changes over time. These analyses are especially useful in answering the prominent questions about access to technical standard-setting fora and conversation, which are important to establishing concepts of fairness and legitimacy when these standards have values implications. I use a combination of email sender metadata and manual annotation to estimate relative fractions of gender within and across mailing lists and I believe similar techniques can (and are, and will) be used to evaluate disparities in participation by sector of affiliation and region of origin.

Network analysis capabilities allow for testing of hypotheses of network formation, for example. Benthall determined that open source development communities did not show the kind of preferential attachment model – a "rich get richer" form of social network development – that has been observed in several other formations of links, as in between websites (2015). Where mailing lists provide important settings for group communication and social network development,

we can use this functionality to measure and compare macro-level properties of these groups. Network analysis can be useful in identifying individuals who play leadership or connecting roles around particular topics or between particular groups.

Content analysis differs from these other types in actually looking "inside the envelope" at the text contents of email messages.[16] Measuring how often words arise can help us see trends in where and how concepts like privacy and security are being discussed. And connecting content analysis with network, demographic or traffic analysis can provide evidence of who is bringing up particular values, how concepts migrate across different groups and when topics see more or less attention.

BigBang developers use a mailing list for discussion, and git and GitHub for sharing source code and coordinating work; it has gone through debates on intellectual property and licensing very familiar to open source software and Internet standard-setting. Use of the tools under study and borrowing a working model similar to the communities under study seems fitting, a scientific analog of "recursive publics" (Kelty 2008).

4.6   Ethics

This work examines people (participants in technical standard-setting, non-participant stakeholders), processes (multistakeholder fora; government, civil society and corporate decision-making; software engineering practice) and architecture (the Internet and World Wide Web). Ethical considerations guide how I have conducted research at each of these scales including in how I direct my inquiry, in protecting human subjects and in handling publicly accessible data.

4.6.1   Studying up   The literature of anthropology has in the past called for more "studying up" — investigating those groups in society that are rich or politically powerful and not, as had been a trend, focusing study on communities that are vulnerable, historically marginalized or foreign in the sense of being from non-Western cultures (Nader 1972). The motivations for this shift are both ethical and scientific; ethical in the sense of not overly objectifying and limiting the scope of

---

[16]While I have explored some use of content analysis for detailing how terminology is getting used or spreading across mailing lists over time, this dissertation does not report on that work or make use of content analysis.

inquiry of problems to those most vulnerable; scientific in the sense of not missing an entire part of society as an object of study.

The ethical impulse to study the culture of people who wield power, and specifically that power held by technical expertise and exercised through the design of influential and immovable technical artifacts, is an essential motivation for this work. The architecture of the Internet and the Web, like many software constructs, have implications for fundamental human values (Nissenbaum 1998); those seemingly technical design decisions are inherently political (Winner 1980), and, like the highway overpasses of Long Island that cast in concrete the impossibility for public buses to reach recreational beaches, have simultaneously long-lasting but hard-to-see impacts (Caro 1975); the processes used for protocol design decisions have opportunities for governance, but also serious open questions for procedural and substantive legitimacy (Doty and Mulligan 2013).

In the study of science and technology, studying up may mean studying the designers: software engineers, developers, user interface designers and "makers" of all kinds, as opposed to studying the larger mass of users whom we typically identify as having less control over these powerful technical decisions. And designers of Internet protocols and Web standards may also constitute an "elite" (Marcus 1983): with agency, some exclusivity and power as their decisions and market status will often influence the decisions of other engineers and technology companies who build on the Internet and the Web.

4.6.2 Interviewing human subjects   To the extent that this research studies individual participants and stakeholders, it qualifies as human subjects research. In particular, semi-structured interviews conducted with standard-setting participants and non-participant stakeholders are aimed not only at understanding the implications of the standard-setting process but at learning about the perspectives, backgrounds and motivations of those individuals.

These conversations are kept confidential to encourage candor from the interviewees and to limit any personal or professional harm that could come from disclosing details of those individuals' perspectives or participation. The names and organizational affiliations of participants is typically not disclosed in this work; quotations are provided with context about the type of participant or organization, but do not include details that could be used to directly identify an individual or their affiliation. Some participants were willing to have quotes directly attributed to them and provided specific consent for that point; their quotes are attributed where I conclude that it provides useful context to the reader.

A research protocol for conducting these semi-structured interviews was reviewed by UC Berkeley's Committee for the Protection of Human Subjects (CPHS, the local Institutional Review Board) and was considered exempt from further review[17] because of the minimal risk of harm in conducting confidential interviews with a non-vulnerable population.[18] After updating with a different funding source and project title, the protocol underwent a much more intensive review process, which resulted in longer (but not more informative) consent forms, updated practices for encrypting data at rest (which were applied to all previous interviews as well), and longer data retention requirements (based on interpretation of federal funding guidelines which I believe to be mistaken). That protocol was approved after "Expedited" review.[19]

4.6.3   Mailing lists   Mailing list analysis also includes collecting and analyzing the communications of human subjects. For this project, list analysis is used both to study the participants and to study the processes and tools of these groups. Because these mailing list archives are collected and publicly presented for the purpose of review, and participants are typically directly informed of this before sending a message to such public lists, Berkeley's IRB provided guidance that no human subjects research review is necessary for this collection and analysis. This is not reducible to an argument that no publicly-accessible data can have any ethical implications for research (an argument which I do not support); these archives are made publicly available specifically for the purpose of access and review by others, including non-participants, and that status is typically well-understood by participants.[20]

Out of politeness, mailing list crawlers were configured to access list archives at no more than 1 request per second. Local analysis of list archives requires making local copies of all the messages of those public archives; those copies may be shared with other researchers in machine-readable form. While those copies are typically identical to the publicly-available archives, we are not currently making those archive copies themselves publicly available. In some rare instances, the hosts of

---

[17]For UC Berkeley's CPHS, "Exempt" does not mean that no review took place or that no review is necessary, but that a research protocol as described contains minimal risk and so does not need to be reviewed by a full IRB committee or yearly updates/reviews.

[18]CPHS Protocol #2011-11-3796.

[19]CPHS Protocol #2018-03-10819.

[20]W3C mailing lists, for example, send an automated reply to any new sender to a publicly-archived mailing list explaining that status and require an explicit form approval by the user before a message is distributed to the list.

mailing list archives will remove messages from the archives even after they have been publicly distributed.

References

Benthall, Sebastian. 2015. "Testing Generative Models of Online Collaboration with BigBang." In *Proceedings of the 14th Python in Science Conference*. https://conference.scipy.org/proceedings/scipy2015/pdfs/sebastian%7B_%7Dbenthall.pdf.

Burrell, Jenna. 2009. "The Field Site as a Network: A Strategy for Locating Ethnographic Research." *Field Methods* 21 (2): 181–99. https://doi.org/10.1177/1525822X08329699.

Cargill, Carl F. 1989. *Information Technology Standardization: Theory, Process, and Organizations*. Newton, MA, USA: Digital Press.

Caro, Robert A. 1975. *The power broker: Robert Moses and the fall of New York*. New York: Vintage Books. http://www.amazon.com/The-Power-Broker-Robert-Moses/dp/0394720245.

Coleman, E. Gabriella. 2012. *Coding Freedom: The Ethics and Aesthetics of Hacking*. Princeton University Press. http://www.amazon.com/dp/0691144613.

Davies, Charlotte Aull. 2012. *Reflexive Ethnography: A Guide to Researching Selves and Others*. Routledge.

Doty, Nick. 2015. "Reviewing for Privacy in Internet and Web Standard-Setting." In *Security and Privacy Workshops (SPW), 2015 IEEE*, 185–92. IEEE. https://npdoty.name/privacy-reviews/iwpe/.

———. 2018. "Standard-Setting Process and Internet Privacy." presented at the Protocol to the People? Protocol governance and power – from Bitcoin to the Border Gateway Protocol, Turing Institute, London, UK, March 16. https://www.turing.ac.uk/events/protocol-people-protocol-governance-power-bitcoin-border-gateway-protocol/.

———. 2019. "Mitigating Browser Fingerprinting in Web Specifications." Interest Group Note. Privacy Interest Group (PING). World Wide Web Consortium. https://www.w3.org/TR/2019/NOTE-fingerprinting-guidance-20190328/.

Doty, Nick, and Deirdre K. Mulligan. 2013. "Internet Multistakeholder Processes and Techno-Policy Standards: Initial Reflections on Privacy at the World Wide Web Consortium." *Journal on Telecommunications and High Technology Law* 11. http://www.jthtl.org/content/articles/V11I1/JTHTLv11i1_MulliganDoty.PDF.

Doty, Nick, Heather West, Justin Brookman, Sean Harvey, and Erica Newland. 2019. "Tracking Compliance and Scope." Working Group Note. Tracking Protection Working Group. World Wide Web Consortium. `https://www.w3.org/TR/tracking-compliance/`.

Geiger, RS, and David Ribes. 2011. "Trace Ethnography: Following Coordination Through Documentary Practices." In *System Sciences (HICSS), 2011 44th . . .*, 0:1–10. Los Alamitos, CA, USA: IEEE Computer Society. `https://doi.org/10.1109/HICSS.2011.455`.

Glaser, Barney G. 1978. *Theoretical Sensitivity: Advances in the Methodology of Grounded Theory*. Sociology Press.

Hennink, Monique M, Bonnie N Kaiser, and Vincent C Marconi. 2017. "Code Saturation Versus Meaning Saturation: How Many Interviews Are Enough?" *Qualitative Health Research* 27 (4): 591–608.

Hine, Christine. 2000. *Virtual Ethnography*. SAGE.

Kelty, Christopher M. 2008. *Two Bits: The Cultural Significance of Free Software*. Duke University Press. `https://twobits.net/`.

Lofland, John, Lyn H. Lofland, David Snow, and Leon Anderson. 2006. *Analyzing Social Settings*. 4th ed. Wadsworth.

Lynch, Michael. 2000. "Against Reflexivity as an Academic Virtue and Source of Privileged Knowledge." *Theory, Culture & Society* 17 (3): 26–54. `https://doi.org/10.1177/02632760022051202`.

Marcus, GE. 1983. *Elites: Ethnographic Issues*. Albuquerque: SAR Press. `http://www.amazon.com/Elites-Ethnographic-American-Research-Advanced/dp/193469133X`.

Massey, Doreen B. 1994. *Space, Place, and Gender*. U of Minnesota Press.

Nader, Laura. 1972. "Up the Anthropologist: Perspectives Gained from Studying Up."

Nissenbaum, Helen. 1998. "Values in the Design of Computer Systems." *Computers and Society* 28 (1): 38–39. `http://www.nyu.edu/projects/nissenbaum/papers/society.pdf`.

Sandelowski, M. 1986. "The problem of rigor in qualitative research." *Advances in nursing science* 8 (3): 27–37. `http://www.ncbi.nlm.nih.gov/pubmed/3083765`.

Strauss, Anselm, and Juliet Corbin. 1990. *Basics of Qualitative Research*. Sage publications.

Wallace, David Foster. 2005. "This Is Water."

Watt, Diane. 2007. "On Becoming a Qualitative Researcher: The Value of Reflexivity." *Qualitative Report* 12 (1): 82–101.

Winner, Langdon. 1980. "Do Artifacts Have Politics?" *Daedalus* 109 (1): 121–36. `http://www.jstor.org/stable/20024652`.