Data ideologies of an interested public: A study of grassroots open government data intermediaries

Andrew Schrock and Gwen Shaffer

Abstract
Government officials claim open data can improve internal and external communication and collaboration. These promises hinge on “data intermediaries”: extra-institutional actors that obtain, use, and translate data for the public. However, we know little about why these individuals might regard open data as a site of civic participation. In response, we draw on Ilana Gershon to conceptualize culturally situated and socially constructed perspectives on data, or “data ideologies.” This study employs mixed methodologies to examine why members of the public hold particular data ideologies and how they vary. In late 2015 the authors engaged the public through a commission in a diverse city of approximately 500,000. Qualitative data was collected from three public focus groups with residents. Simultaneously, we obtained quantitative data from surveys. Participants’ data ideologies varied based on how they perceived data to be useful for collaboration, tasks, and translations. Bucking the “geek” stereotype, only a minority of those surveyed (20%) were professional software developers or engineers. Although only a nascent movement, we argue open data intermediaries have important roles to play in a new political landscape.

Keywords
Community, data intermediaries, data workers, local government, open data, politics

Government officials and policy-makers have high hopes for “open data.” Barack Obama’s 2009 memorandum Transparency and Open Government positioned open government as transparent, participatory, and collaborative. He defined the default medium at the federal level as “open and machine-readable” data, rhetorically framing open data as a form of open government. Governments interested in open data make data available to the public under an open-source license. Officials hope open data leads to positive internal outcomes for cities such as measuring programs and improving efficiencies (Kitchin, 2014: 56). Open data is thought to be useful to involve citizens more directly in decision-making processes and enable collaboration on technological projects (Yu and Robinson, 2012). Civil society groups and journalists hope to make use of open data to empower distributed publics (Baack, 2015). In these ways backers of open data hope to improve government’s internal and external communication.

The core ideas behind open government data in the United States emerged from the right to information movement and the signing of the Freedom of Information Act in 1965 (Schudson, 2015). In 2007 a group of tech and policy luminaries including Carl Malamud, Tim O’Reilly, a young Aaron Swartz, and Lawrence Lessig met at Sebastopol, California (Schrock, 2016). The open data definition that emerged focused on eight qualities of data: completeness, priority, timeliness, ease of physical and electronic access, machine readability, non-discrimination, use of commonly owned standards, licensing, permanence, and usage costs. While the situation regarding open
government data varies significantly internationally (Davies, 2014), in the United States local open data initiatives often draw from Sebastopol and Obama’s federal-level definitions, which define open government data as facilitating transparency, participation, and collaboration.

Surveys show the public is generally receptive to the idea of open government data but have vague ideas about how it relates to their lives. A nationally representative survey administered by the Pew Center for Internet & American Life found the public rarely associate open government data with collaborative definitions (Horrigan and Rainie, 2015). That is, the public has little understanding of how to access and interpret data sets. Accordingly, researchers have generally focused on “open government data intermediaries” that download, interpret, and manipulate data (Dumpawar, 2015; Magalhaes et al., 2013; Sawicki and Craig, 1996). These extra-institutional data experts are vital in our current era when government uses data to make decisions that impact the public good. Yet, little is known about how open government data intermediaries conceptualize the civic benefits of data and are able to connect it with particular uses beneficial for their community (Gurstein, 2011).

To understand the organizational potential for open data to improve democratic institutions requires being attentive to how intermediaries connect data with particular civic goals. Communication scholars interested in civic engagement or politics have mostly gravitated towards describing why intrinsic properties of a particular technology alter participation. Technologies—predominantly the internet or social media—alter the form and function of political action. For example, Bennett and Segerberg (2012) suggested that new organizational logics emerge around “personalized content sharing across media networks” (p. 739). Social media changed the question of participation from collective action to “connective action.” Our concern is that taking a social shaping perspective draws attention away from how individuals develop understandings about the political utility of particular technologies. Being attentive to how effective uses are socially constructed can reveal alignments and obstacles to organizing and acting politically (Dunbar-Hester, 2014). This study is, in part, a test of the claim of government officials that open data naturally leads to seamless collaboration. We also hope to add to the depth of existing studies on open government data that look at use of platforms and potential outcomes as an indication of civic value.

This article proceeds by outlining what is known about the motivations of open data intermediaries. We were inspired by Couldry and Powell’s (2014) call to foreground “the agency and reflexivity of individual actors as well as the variable ways in which power and participation are constructed and enacted” (p. 1). It then draws on Ilana Gershon’s notion of “media ideologies” (Gershon, 2012, 2010) to conceptualize “data ideologies”—culturally and socially inscribed beliefs about the appropriateness of data for certain communicative purposes. Finally, we used qualitative (focus groups) and quantitative (surveys) methodologies to construct how data ideologies varied across different groups. This mixed-methods approach resulted in depth of understanding of where data ideologies come from, and a basic understanding of how they manifested in a particular community in Southern California.¹

Open data intermediaries

“Open government data intermediaries” are defined as actors that translate, use, or otherwise mediate communication using data produced by or for government (Magalhaes et al., 2013). The term “data intermediary” was coined in 1996 by researcher-practitioners Sawicki and Craig (1996). They believed certain individuals served a “democratizing” purpose because they provided “access to data and assistance with analysis and policy development” (p. 512) to community groups. Data intermediaries made information available and distributed it to interested users (Sein and Furuholt, 2012). The “open government data intermediary” is a particular genre of data intermediary: non-state actors who work with data in a larger ecosystem involving government actors, companies, and citizens. Open government data intermediaries have been found to include civic startups, open data services, and infomediaries (Magalhaes et al., 2013). Most importantly, open government data intermediaries make use of data in novel ways, connecting entities that would not otherwise be connected. Van Schalkwyk et al. (2015) summarize that open data intermediaries increase the utility of data, move it throughout a data ecosystem, and serve a democratizing function.

Significant barriers exist to open government data intermediaries achieving civic goals that benefit the public. Kitchin (2013) noted several barriers: increasing participation and awareness of the politics of data; connecting with effective uses; and moving away from purely corporate applications. In his successive book *The Data Revolution*, Kitchin further argued for paying greater attention to sustainability and “how data are being utilized and employed” (p. 64), rather than purely technical measures of success such as website hits and data sets downloaded. Conversely, Gurstein (2011) pushed to consider barriers to “effective uses” of data. Only by connecting data with
actors and their interests might the “virtuous cycle” of open data lead more easily to collaborations between government, civil society groups, and residents. The primary issues confronting open data intermediaries are discussed in turn.

Access is often used as the measure of whether open government data is effective. However, as Kitchin (2014) noted, data are never neutral or prima facie lead to positive outcomes. There are even reasons why even available data may paradoxically distance governments from communicating with citizens. Andrejevic (2014) described a “Big Data divide”—a new type of asymmetrical power relationship between individuals who generate data and those who put it to use. He noted this is a question of access, technical skills, and also the infrastructure to make use of data. If an individual downloads their data from Facebook, they still remain unable to match advertisements to it and generate revenue. If governments privilege open data as a mode of civic engagement over other forms, those unable to participate will simply be left out (Gurstein, 2011). Johnson (2014) similarly notes that a focus on access tends to downplay how data often carry assumptions of privilege and data users might have limited technical capabilities.

Open data intermediaries may also not share the same understandings of political participation. Policy-makers and urban planners have particularly emphasized the participatory dimensions of open data (Goldstein and Dyson, 2013; Townsend, 2013). Open data intermediaries are a natural way to achieve goals of collaborating with the public with limited government overhead. Yet, open data intermediaries might rarely have the same goals. As Emily Shaw (2014) noted, “what motivates governments to create an open data initiative may differ from the goals that citizens’ groups, community service organizations, journalists, academics, or civic hackers may have in using open data” (p. 121). Suruchi Dumpawar (2015) suggested that data intermediaries use data for a wide variety of purposes such as aggregation, rectification, interpretation, representation, dissemination, augmentation, and connection.

The various groups that organize around open data may operate as a barrier to acting in unison. While a complete review of effective uses of open government data is outside the scope of this article, three types of open government data intermediaries have been the focus of scholarly attention: activists (Baack, 2015), journalists (Schrock, 2016), and hackers (Kubitschko, 2015). These groups reflect quite different norms about how open government data should be used. Baack (2015) makes a case for activists as extra-institutional actors that use open data for political purposes. According to Baack, they believe that “sharing raw data should help citizens to better understand their governments and to be more active and engaged in their local communities” (p. 5). Data journalists use data to create multimedia stories that put pressure on governments and corporations to undertake reform efforts (Parasie and Dagiral, 2012). Finally, hackers can assist the public good by using data and software literacies to improve their local community and government. For example, the Chaos Computer Club (CCC) in Germany translated technical issues for the public and assisted in the crafting of government policy. Kubitschko (2015) termed this relationship of a hacker group to institutional concerns an “interlocking arrangement.”

In summary, this article clarifies how non-state actors act as open government data intermediaries in the context of a city of approximately 500,000 in Southern California. Although intermediaries have been of scholarly interest, few have theorized why open government data intermediaries hold the beliefs they do about the civic benefits of data. This article clarifies this ambiguity by applying theory about how they conceptualize civic participation and effective uses. Specifically, we conceptualize that “data ideologies,” as a type of media ideology (Gershon, 2010, 2012), drive the use of open government data.

Data ideologies

Individuals hold particular beliefs about data that are socially inscribed and culturally informed by their upbringing, schooling, and participation in communities of practice. For example, Fiore-Gartland and Neff (2015) coined the term “data valences,” to refer to “people’s expectations of and values for data that emerge from their discourses and practices across different contexts.” Open government data is of particular interest because, as described above, it is often ideologically imbued with particular values and inspires practices that benefit the public. We were therefore motivated to take an approach that was more attentive to how individuals envisioned themselves working with data.

In this paper we draw on Ilana Gershon (Gershon, 2010, 2012) to define “data ideologies.” Gershon’s concept of media ideologies refers to beliefs people hold about the appropriateness of a medium for a particular communicative goal. Gershon was keenly interested in the motivations behind use of a technology for breaking up with a significant other. Media ideologies’ theoretical framework emerged from linguistics, particularly semiotic ideologies. Gershon believed there was little innate about a particular technology that makes it more appropriate for a particular task. Rather, individuals’ perception of appropriateness of technology for particular purposes was socially and culturally constructed.
Certain ideas were points of differentiation in acceptable uses for particular technologies. Formality (Gershon, 2012: 23) was one way that individuals decided on whether a particular medium was appropriate for a particular goal. Media ideologies are “always multiple, locatable, positioned, and contested” (Gershon, 2012: 33). One person might deem texting too informal to be appropriate for ending a relationship, while another person might find it suitable. Key terms such as formality—which emerged from her qualitative research—were distinctions between the ways individuals thought about and used media for particular purposes.

Two concepts were particularly central to Gershon’s thinking: remediation and idioms of practice. Remediation (Bolter and Grusin, 2000) described how “new media” are rarely a dramatic break from previous media. People understand how new communication technologies are used in relation to earlier forms. For example, young individuals might compare email to the formality of sending a letter. The enthusiasm for attaching civic outcomes to open data demonstrates how it captures a certain political valence by remediating previous forms of openness (Schrock, 2016).

Gershon believed that shared understandings of appropriate practices with a particular medium emerged through communication in particular groups, or “idioms of practice” (Gershon, 2012: 39). Communicating in social and professional groups permits ways for mistakes to be made that break social norms. That is, the acceptability of a medium for specific purposes is established only through shared experimentation, shared engagement, and observation. New technologies only coalesce into widely accepted practices—what we might call common or normative uses—over long periods of time (Gitelman, 2004). During early years of a medium’s introduction people can have widely varied ideas about acceptable uses.

Data can be approached as a medium because it is inscribed with meanings, transmitted, decoded, and interpreted through specific practices (Hall, 2006). Gershon’s notion of media ideologies is useful for exploring how individuals espouse subjective beliefs about the utility of data for civic purposes. It bears mention that a media ideologies approach bears similarities to sociological kin focusing on the cultural and social shaping of practices. Cognitive sociology has long considered how cultures and occupations develop particular ways of “seeing” particular objects (Zerubavel, 1999). To Zerubavel, shared engagement with particular subcultures teaches us how to perceive and interpret the world. Of these theories, we selected media ideologies because of its history in communication research, social shaping perspective, and strong alignment with action.

Similar to Gershon, we suggest that there is little inherently “civic” about open data. The civic-ness of open data—how it improves shared life circumstances in various ways—is socially constructed. People regard norms of civic participation differently across history (Schudson, 1999). Open data is a relatively new concept and people are still figuring out how it remediates older forms of civic action or creates new potentialities. Although we find media ideologies an appropriate theoretical base, our medium of choice and research questions directed us to conduct research in a slightly different way. Gershon was concerned about the way individuals negotiate a single common interpersonal action (breakups in romantic relationships) across multiple media. Often this communication was in public or semi-public environments such as Facebook (Gershon, 2012: 165). By comparison, we were interested in examining how beliefs about differences in how data (a single medium) was related to civic action.

Methodology and results

The current study employed public participation in a local government’s open data initiative to research the prevalence and character of data intermediaries. Local government is a promising site for improving citizen participation in such initiatives. Local government tends to emphasize autonomy, participation, and efficiency (Kjellberg, 1995). It is more trusted by residents (63%) than federal (28%) or state (57%) levels (Dimock et al., 2013). As Christopher Ansell observed, “public agencies... are the designated leads for concrete problem-solving” (Ansell, 2011: 183). According to a 2015 survey of city governments in Southern California, a majority of city governments (76%) said open data is a moderate to top priority, and 41% agreed that it “empowers citizens to solve city problems” (USC CT, 2015).

To surface data ideologies for analysis, we used a novel sequence of quantitative and qualitative methodologies to collect and analyze data (Figure 1). We used public focus groups for a local open data initiative as an opportunity to collect and organize ways people think about civic uses for open data. A city of around 500,000 in the southwestern United States tasked its Technology & Innovation Commission with drafting an open data policy. The authors, as two members of this commission, collected empirical evidence about how an interested public views open data. These perspectives were then integrated into the open data policy. There is a strong need for resident outreach and customization. A recent survey that included most municipal governments in Los Angeles county concluded that cities may well need to “go beyond existing open data criteria to make sure open data is
working in the unique context of each city” (USC CT, 2015: 23). Crafting an open data policy therefore requires addressing who might serve as data intermediaries for this particular community.

We took a grounded approach (Charmaz, 2005; Glaser and Strauss, 1967) for analysis, allowing motivations for using data to arise organically from participants’ own words. Our research emerged from shared participation as members of a Technology and Innovation (T&I) Commission in an ethnically, racially, and economically diverse city of 500,000 in southern California. The Director of T&I was interested, as were we, in the city’s open data policy reflecting the needs of the public. In response the T&I Commission organized three public focus groups designed to foster discussion between officials and a public about concerns and potential uses for data.

“Public” refers to the notion of a public activated by what Dewey termed “shared consequences” (Dewey, 1927). Dewey draws our attention to individuals’ capacity for empathy and forethought around particular issues (Marres, 2007). To Dewey, a public “consists of all those who are affected by the indirect consequences of transactions, to such an extent that it is deemed necessary to have those consequences systematically cared for.” Government officials and advisory groups are necessary to cohere publics around issues of public concern and to help instruct them how to act. This is different from a quantitative social science perspective, which attempts to obtain a sampling that is statistically representative of a particular population. The prevalence of open data intermediaries is unknown. However, the literature shows they are fairly rare and hold a variety of different perspectives on how to use data, making a representative sampling unhelpful for obtaining more textured understanding of their perspectives. For our research, the open data focus groups were publicized through online message boards maintained by the government, flyers, and word-of-mouth. There was no guarantee we would be able to have an equal probability of reaching each resident. We therefore oriented our sampling around a Deweyian notion of public rather than a representative sampling, as is more common in society science.

The three focus groups were held at a downtown business, a community college, and a university during various times of the day. A total of 73 participants were counted at the focus groups. The sites were geographically dispersed and easily accessible by mass transit. The researchers distributed surveys at each event prior to discussions about open data. An identical version of the survey was posted online for participants who preferred an electronic format, as well as for residents who did not attend the forums. The city promoted the survey through its official social media platforms and website. A total of 70 surveys were collected, including online and offline collections.

This study used a multi-step process to obtain a nuanced and accurate perspective on how an interested public in the city viewed the utility of open data. The first two steps used qualitative methodologies of focus groups and participant-observation by researchers. In the round-table setting we asked participants standardized questions, such as “how would you use open data?” During each session we collected details about the discussion, and reflexive notes to ourselves about the ways our preconceptions about data workers and intermediaries were changing. After each forum we assembled user personas (Larman, 2003) to represent specific participants. User personas are a technique from design to capture individuals’ motivations and perspectives using particular technologies. They consist of a brief description of their personal history and goals, much like a short biography. However, unlike designers, we sought to keep personas true to the individuals. Our personas served to retain detail and accuracy as we refined our notes and move towards thinking about particular idioms of practice. At this point in the study, the authors were primarily concerned with recording participant data. A total of 21 personas were created, one of which is shown in online Appendix.

The second step was refinement, open coding, and axial coding to organize variation in data ideologies. We used open coding to analyze the persons in combination with qualitative data (notes and personas) for similarity and difference (Berg, 2004). Finally we created categories that captured individuals’ interests and beliefs about how open data should be used. This stage of the process was similar to open coding in grounded theory (Glaser and Strauss, 1967) where clusters of similar themes are created and subjected to axial coding. The resulting typology—collaboration, tasks, and translations—is discussed at length in the next section of the paper. The final step in the methodology was to collect quantitative statistics on surveys for demographic (Table 1) and attitudinal (Table 2) variation. Surveys provide categorical and ordinal level data that reveal information that are difficult to collect in a small group setting. This step revealed important information about individual factors impeding the crossing of the data divide that were not visible from the open data forums. For example, while a participants’ occupation often came up in discussion, educational level and age typically did not.

**Distinctions in open data ideologies**

This article suggests a provisional typology that describes categorical variation in data ideologies among open data intermediaries. Gershon used
qualitative methodologies to formulate vocabulary that
described how individuals interpreted the appropriate-
ness of media for relational dissolution. Similarly, here
we suggest a set of distinctions that emerged across
participants in the focus groups. Similar to Gershon,
these concepts do not describe universal perspectives
but rather points of disagreement about how a
medium should be used. We were being attentive to
disputes, or as Howard Becker put it, “looking for
trouble” (Becker, 2007: 129). Our typology describes
distinctions in ways participants regarded the appropri-
ateness of data for civic purposes.

**Collaboration**

Making data freely available to re-use was fundamental
to participants’ conception of “open data.”
Participants espoused a belief that it should be used
to improve their local government and community
life. However, exactly how open data would bring
about government change was hotly debated. The
most palpable distinction was whether releasing open
data would put pressure on government to act, or if
open data intermediaries would be government colla-
borators. These two perspectives follow in Jane
Mansbridge’s (1980) distinction between adversary
and unitary forms of democracy. Adversary forms of
democracy are based on conflicting interests, equal
protection, majority rule and secret ballot (p. 5). Unitary
democracy is based on equal status, consensus,
common interest, and (often) face-to-face contact in
collaborative meetings. Followers of unitary democracy
believe they are “all in it together,” while adversary
democracy advocates believe they need to “have their
voice heard.”

Margaret, a local activist, took an adversarial
approach to open data. She was leery of the promises
of a seamless ecosystem of data interpretation and use.
Dave, a professional journalist, similarly lamented
“we’re always in a battle for data.” To Dave, mean-
ingful structural change—government changing its
ways—came only when data was actively requested
and revealed facets of internal operations that govern-
ment would rather suppress. Getting data was part of a
give-and-take between an active citizenry and govern-
ment officials doing nefarious things. Moments of
public scandal that activists and journalists brought
to light were productive to shame officials into more
responsible behavior. Members of the local coding
and tech community saw demand for “raw data” as
leading to government accountability, or what a
dubious Lawrence Lessig termed “naked transparency”
(Lessig, 2009). Providing open data would ensure
accountability because officials would understand that
their actions would be subject to scrutiny and alter their
behavior. According to Matt, a technically proficient
local business-owner, believed the city should adopt
aspirational language that associated open data with
transparency. He believed that the release of “raw
data” served a more practical purpose, to improve com-
munication practices: “we don’t have to beg for data.”

Those who believed in a more unitary form of democ-
Jacy saw open data as letting them act as external advisors. Sam, a student and employee at a local “big box”
store working in transportation (online Appendix),

### Table 1. Demographic variation in open data survey participants.

<table>
<thead>
<tr>
<th>Gender (N = 63)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>62%</td>
</tr>
<tr>
<td>Female</td>
<td>37%</td>
</tr>
<tr>
<td>Other</td>
<td>2%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Race/Ethnicity (N = 62)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>White or Caucasian</td>
<td>42%</td>
</tr>
<tr>
<td>Black or African-American</td>
<td>5%</td>
</tr>
<tr>
<td>Asian</td>
<td>16%</td>
</tr>
<tr>
<td>Latin-American Ethnicity/Hispanic</td>
<td>23%</td>
</tr>
<tr>
<td>Multiracial</td>
<td>5%</td>
</tr>
<tr>
<td>Other</td>
<td>10%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Education (N = 66)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>High school or lower</td>
<td>6%</td>
</tr>
<tr>
<td>Some college, did not graduate</td>
<td>29%</td>
</tr>
<tr>
<td>College (2-year or 4-year degree)</td>
<td>45%</td>
</tr>
<tr>
<td>Graduate school (MA or PhD)</td>
<td>20%</td>
</tr>
</tbody>
</table>

### Table 2. Skill and Attitudinal variation in open data survey participants.

<table>
<thead>
<tr>
<th>Computer experience (N = 67)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Beginner</td>
<td>8%</td>
</tr>
<tr>
<td>Average</td>
<td>24%</td>
</tr>
<tr>
<td>Advanced</td>
<td>39%</td>
</tr>
<tr>
<td>Expert</td>
<td>30%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Familiarity with open data (N = 64)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Not familiar</td>
<td>28%</td>
</tr>
<tr>
<td>(between not and somewhat)</td>
<td>9%</td>
</tr>
<tr>
<td>Somewhat familiar</td>
<td>28%</td>
</tr>
<tr>
<td>(between somewhat and very)</td>
<td>19%</td>
</tr>
<tr>
<td>Very familiar</td>
<td>16%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Government trust (N = 67)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>None of the time</td>
<td>1%</td>
</tr>
<tr>
<td>(between none and some)</td>
<td>24%</td>
</tr>
<tr>
<td>Some of the time</td>
<td>48%</td>
</tr>
<tr>
<td>(between some and all)</td>
<td>25%</td>
</tr>
<tr>
<td>All of the time</td>
<td>1%</td>
</tr>
</tbody>
</table>
envisioned an ecosystem where he could obtain environmental data that he could interpret. Eventually it would be delivered back to modify the government entities that produced it to help them better understand how to act. For example, to retain the city’s claim of a “green port” experts needed to perform more detailed analyses of pollution. Sam imagined he would help interpret data to help receptive officials make wiser decisions about transportation infrastructure. The Chief Innovation Officer similarly espoused beliefs that “open data” should be used in collaborative ways that never caused embarrassment and was always a sensible return on investment. Citizens could “pitch in” by working with data in ways the city had little time or interest in. The city’s moderate size was a point of contrast to larger cities like New York City and Philadelphia with burgeoning Civic Tech scenes. Local government did not have infinite resources, so the collaborative message of open data was appealing for financial reasons, even if they didn’t have the labor to just make everything available in the way Matt imagined.

**Tasks**

Tasks learned in professional occupations were a recurring theme underlying many perspectives. Participants learned technical skills and developed ethical convictions about data through their jobs. Returning to Gershon’s concept of “idioms of practice,” communities of practice remediated previous practices with data. For example, Dave, a professional journalist, saw open data as holding the potential for new stories. Journalists have a long history in obtaining, translating, and creating knowledge that serves the public good from data (Schudson, 2015). Open data was a way for him to do his job, which had an established set of ethical convictions and practices that aligned with a cultural “right to know.” Others were very familiar with other forms of data through their work but didn’t see how it could be used in the context of open data. Graham worked extensively with geographic data and built multimedia tours for mobile devices in areas beyond the reach of cellular connectivity. Despite his expertise, he was only interested in opening up financial data, leading to the kind of naked transparency espoused by civic tech enthusiasts like Matt and Robert. Graham saw not just different types of data; he saw data as imbued with particular political valences. Of course, many in the open-source mapping community regard geographic data as a site of political struggle. Creating data sets through crowdsourced methods that are then released to the public is a way to encourage growth of community-based mapping systems and empowering individuals.

Participants from the corporate world believed that open data could be useful to tracking and “nudging” residents to pro-civic behavior. Julian, an energetic student from a local University, saw open data from a marketer’s perspective. He had minimal exposure to the idea of open data, so he applied what he learned about data from his undergraduate courses. He thought of open data as including behavioral traces (similar to Facebook and Google) that could be used for selling advertisements. James, a government official for the city tasked with economic revitalization, was interested in how open data could serve as an informational resource for local businesses. In his mind, a local business owner (say a pizza shop) might need information about licensing. They go to the open data portal and obtain information that helps them make decisions. He believed that data moving from inside to outside government could be a resource in much the same way as any informational website. He also saw businesses as
being able to launch enterprises off open data, creating social media platforms much like FourSquare. Open data was the next iteration of government’s evolution from static websites to being more interactive and responsive. Open data became a vehicle for economic development.

**Translations**

Open data was nearly universally regarded as cost-neutral “exhaust” from government operations that could be converted. Its mutuality led to speculation about the potential for creating mobile apps, statistics, visualizations, and stories. Why and for whom they translated data was the topic of vigor debate. Some sought to mobilize the local community to take action on particular issues. Gerard, who had little formal training, taught himself how to obtain and interpret data related to speeding around schools. After witnessing a near-fatal accident, he was trying to gather support for policy changes among policy-makers and community members. He only reluctantly used the term “activist,” which he thought was too politically charged. He sought to mobilize the community by translating data into information that he relayed on social media and flyers. Sam was interested in interpreting the meaning of arcane scientific and transportation terms. He gave an example of how an agency might provide data on the number of overall particulates but not their relative size; a larger particulate might be more harmful to individuals than an equal amount of smaller particulates.

Gerard and Sam created stories, visuals, and other informational media to be relayed online, in pamphlets, and publications. Both were keenly aware that data was constructed and may not represent objective scientific facts. Matt, by contrast, was more interested in seamless utility—integrating data into data-driven mobile applications (“apps”). A few weeks prior to the forum equipment failures in the electrical grid caused blackouts downtown. Matt used this scenario as an example of the type of issue that open data could help prevent if it was integrated into a mobile application. James, working in local government, saw it both ways; data was both informational and capable of driving new mobile and online applications. Open data was a kind of meta-medium that existed to be transformed to other media that were interpretable by non-tech-literate people.

Finally there were a small number of participants that showed up to the forums who were certainly civic leaders but not open data intermediaries. Anne-Marie was retired and assisting the city by participation in a commission for elderly issues. She was unfamiliar with the topic of open data and saw no way of connecting it with her service. She didn’t have the technical experience or interest in open data. To her, data was a barrier to the coalition-building and advocacy work that was Anne-Marie’s primary mode of interacting with government. Given our recruitment strategy it isn’t entirely surprising that a few people showed up that were more curious than already familiar with open data. Anne-Marie is a valuable example of an “edge case” of someone civically involved that might work with open data intermediaries but not ever become one herself.

**Discussion**

Being attentive to perceptions about data enables us to better understand the civic potential for institutional interventions that orient in part around technology. Data ideologies describe ways open data intermediaries imagine uses of data being mapped onto civic goals. These data ideologies varied by how they conceptualized collaboration, tasks, and translation. Several participants fit into known categories of data intermediaries (such as journalist and activist). Others were new intermediaries in an open data ecosystem, or merely curious, at the beginning of a potential path to civic awareness. Political convictions and community commitments were interwoven throughout, even though some privileged building a cool mobile app over mobilization.

Several limitations of this study exist. The survey and focus groups drew from a convenience sampling of local residents rather than a representative sampling frame. Therefore, conclusions from this study are limited to this particular city and the local level of government. Open data intermediaries active at the state or federal level might exhibit quite different motivations, partly because these levels of government are responsible for different government functions. Future research might consider how to obtain a representative sampling within a particular region such as a state or country. Demographics from surveys show that an interested public was relatively young, white, and educated. One limitation of the current study is it does not capture the experiences of groups on the margins who have negative experiences with the government in particular may maneuver around government systems because they are not comfortable interacting with government officials (Brayne, 2014). On one hand, this paints a picture of engagement through data on issues by the relatively educated and affluent. Yet, contrary to expectations, participants were not only software developers and showed only a moderate degree of trust in government. Many participants were self-taught and brought participatory expertise with geographic and environmental data.
These varied practices of individuals with a shared interest in open data may eventually cohere into a more formalized movement. “Civic tech,” for example, is a term adopted by leftist geeks who seek institutional reform and equality in the way that government interacts with, serves, and represents residents. There is a strong overlap between civic tech and open data intermediaries, and for the last several years civic tech has been on the cusp of forming into a broader movement. Open data enthusiasts and civic technologists have capitalized on the semiotic notions of openness similar to the political mobilization around free and open-source software (Coleman, 2012; Kelty, 2008). Yet, this study indicates that open data intermediaries lack a shared culture and political understandings necessary for broader and more impactful action.

For all their bold ideas about improving community life and government resilience and interest among policy-makers in data-driven collaboration, grassroots participants’ perspectives were exceeding fragmented. Data intermediaries held a variety of ideas about what data is and might do. Gerard’s model of community organizing—using data to translate an issue to mobilize his neighbors—had little in common with Matt’s radical openness or Dave’s principled journalistic storytelling. Gershon reminds us to expect such unsettled practices from a new medium. Perhaps this is especially true of open data, a concept that has a complex and even divergent political history (Schrock, 2016; Schudson, 2015). Domain knowledge required for participation and widely differing goals may make it difficult to act collectively. Christina Dunbar-Hester (2014) observed that technical cultures find it difficult to mobilize because they are doubly-bound by limitations of technological literacies and differing political aspirations.

Recent developments have drawn into question the future of open data intermediaries. Political developments in late 2016 could act as a galvanizing force to bring together these different actors on the political left. The funding issues of the Sunlight Foundation and the presidential nomination of the Donald Trump are the most serious challenges to transparency efforts among open data enthusiasts in the last decade. Open data intermediaries may become more adversarial at the federal level—pushing for disclosure of information along side the press—while continuing to be more unitary locally, collaborating openly with sympathetic democratic mayors and governors to improve government. Future work might also consider the local question of a “data ecosystem” in communities; what are binding potentials for civic action among non-profit groups, collectives, and schools involved in data issues? While the political situation is still uncertain and unfolding, open data intermediaries clearly have important roles to play in this political landscape.

Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

Supplemental Material
Supplementary material for this paper can be found at http://journals.sagepub.com/doi/suppl/10.1177/2053951717690750.

Note
1. The policy implications of our data collection have been mostly retained for a subsequent paper. The current paper focuses on an in-depth analysis of the character and motivations of data intermediaries.

References


