“Wiring a City”: A Sociotechnical Perspective on Deploying Urban Sensor Networks

LUCY VAN KLEUNEN, Department of Computer Science, University of Colorado Boulder, USA
BRIAN MULLER, Environmental Design, University of Colorado Boulder, USA
STEPHEN VOIDA, Department of Information Science, University of Colorado Boulder, USA

We use a sociotechnical perspective to expand upon prior characterizations of deploying end-to-end urban sensor networks that focus primarily on the technical aspects of such systems. Via exploratory, semi-structured interviews with those deploying a number of urban sensor networks in a single American city, we identify ways that human decision-making and collaborative processes influence how these infrastructures are built. We synthesize these findings into a framework in which sociotechnical factors show up across the phases of data collection, management, analysis, and impacts within smart city projects. Each phase can display variability in immediacy, automation, geographic scope, and ownership. Finally, we use our situated work to discuss a generalizable tension within smart city projects between cross-domain data integration and fragmentation and provide implications for CSCW research, the design of smart city data platforms, and municipal policy.

CCS Concepts: • **Human-centered computing → Empirical studies in collaborative and social computing**.

Additional Key Words and Phrases: smart cities; urban sensor networks; sociotechnical system; civic data

ACM Reference Format:
Lucy Van Kleunen, Brian Muller, and Stephen Voida. 2021. “Wiring a City”: A Sociotechnical Perspective on Deploying Urban Sensor Networks. *PACM on Human-Computer Interaction* 5, CSCW1, Article 178 (April 2021), 22 pages. https://doi.org/10.1145/3449252

1 INTRODUCTION

Urban areas around the world are struggling with the overlapping challenges of rapid growth in population and increasing natural disasters due to climate change. Local governments are experimenting with novel sensing technologies to address these challenges. However, a city is already a complex social, technological, and ecological system. In this work, we complicate the notion of “*wiring a city*” (P9) with sensor networks by investigating how new technical infrastructures are integrated into the existing streets and neighborhoods of a particular city.

The dominant vision of the “smart city” that is put forward by technology industry players [23], echoed in national funding streams in the United States [39], and investigated by the technical research community [25] involves instrumenting environmental and social sub-systems (transportation routes, streams, etc.) in cities with data collection infrastructure. The promise is that

Authors’ addresses: Lucy Van Kleunen, lucy.vankleunen@colorado.edu, Department of Computer Science, University of Colorado Boulder, Boulder, Colorado, USA; Brian Muller, brian.h.muller@colorado.edu, Environmental Design, University of Colorado Boulder, Boulder, Colorado, USA; Stephen Voida, svoida@colorado.edu, Department of Information Science, University of Colorado Boulder, Boulder, Colorado, USA.

2573-0142/2021/4-ART178 $15.00
https://doi.org/10.1145/3449252

these technical systems will yield data that can be used to inform city policy, and that combining and analyzing data from these sensor networks in near–real-time can allow for intervention if issues are detected.

Previous work in the technical literature has identified challenges in deploying smart city infrastructure, including the integration of heterogeneous data types across domains—for example, combining data from a traffic monitoring sensor network and an air quality monitoring sensor network [38, 56]. However, many of the technical challenges in smart city deployment also have significant—even dominant—social and organizational components. For example, as demonstrated in a quote by Participant 11 in our study, integrating heterogeneous data types also means setting up collaborations between different city departments as well as with private companies:

> *Public Works or the Transportation Department might know what's up on the street; the Parks Department might be in charge of the trees. So it's like there's a lot of different entities that have interest in the public right-of-way, but there's no like one owner of it all that can like manage it super effectively or efficiently. Meanwhile, we're giving it to the private sector often for very low costs and sometimes for nothing to crowd up that space.* (P11)

There has been considerable recent interest in the politics of big data systems, including in how human biases can be encoded into data infrastructures [5]. In the public domain, urban sensor networks are data-intensive systems that influence policy decisions and thus the lives of city residents. Local government actors, working under societal constraints, make decisions about how and what data is collected, how it is combined and analyzed, and how the information produced closes the loop and impacts the world. It is essential to develop a framework through which these decisions and their rationale can be uncovered to understand, for example, if and how biases might be introduced into smart city systems.

We approached this study with the following research question: *How do sociotechnical factors influence the design and deployment of smart city systems, and thus the architecture and flow of data within such systems?* We collected the perspectives of those deploying smart city infrastructure within a specific American city. Our focus was on large-scale projects that involved sensors and in which a local government was a dominant stakeholder. We did not focus on public perception of smart city technology except insofar as those we interviewed attempted to anticipate public response during the deployment and maintenance of urban sensor networks.

Our study reveals how the process of collecting, storing, analyzing, and using data from urban sensors is mediated throughout by people and organizations and fundamentally influenced by geography and politics. We contribute a rich description of a number of overlapping smart city projects in one American city, along with an emergent framework for supporting future analytic work in this domain.

Our analytical framework traces the flow of data through the phases of data *collection*, *management*, *analysis*, and *impacts*. Each of these phases can display variability in *immediacy*, *automation*, *geography*, and institutional *ownership*. For example, a smart city project might have automated, immediate data collection infrastructure in a single city that is owned by a private company, data might then be managed in a publicly owned state database and automatically linked to other data sources, it might be analyzed by human analysts over the course of months, and then may impact the world by both being displayed on an online public dashboard and used to design educational programs. In a purely technical view of smart city infrastructure, the human elements of this system (people physically carrying samples, performing analyses, or educating others about data) might be missed, as would the variations in geographic scope due to political geography and ownership due to public–private partnerships. By providing this narrative and framework, we expand upon prior characterizations of end-to-end urban sensor networks that focus more narrowly on the technical aspects of such systems (see [17, 38, 56]) or more exclusively on organizational and political factors (e.g., [27]).

Finally, we also reflect upon how our situated work helps in understanding a generalizable tension in smart cities between cross-domain data integration and fragmentation. While the smart city vision often idealizes a “system of systems” paradigm [38] in which data from different domains is combined into a common hub and analyzed together, our work revealed a mix of technical, social, and organizational tensions that work to reduce the amount of actual cross-domain data integration happening in real world smart cities. Our work opens the door for further investigation of urban infrastructure development and operation as a site well-positioned for future study in the domain of computer-supported cooperative work (CSCW).

2 RELATED WORK

2.1 Smart cities and urban sensor networks

The precise definition of a “smart city” has been widely characterized as ambiguous within both computing and policy literature. Smart city-branded programs undertaken by cities can include a wide variety of projects and services that involve the use of information and communications technologies (ICT) in an urban context, as well as a variety of sensing strategies [17, 40, 55]. In this work, we focus on urban sensor networks—what Neirotti et al. would characterize as the “hard” smart city approach rather than the “soft” approach, which focuses on e-government and promoting innovation [40]. We chose this focus in order to reduce ambiguity in our analysis, and because the “hard” vision of the smart city most closely aligns with how national funding sources, private companies, and individual cities in the United States describe smart city projects [9, 25, 39, 40]. Although we scoped our investigation to urban sensor networks, we also asked participants to reflect specifically on the more-loaded “smart city” term and how it did or did not accurately describe their work.

2.2 Critiques of smart cities

Those critical of the smart city concept have argued that they may prioritize inappropriate ideals from computer science, such as ubiquity and invisibility [19], and thus help cities position themselves as innovative without actually addressing inequality or other core challenges related to urbanization [2, 64]. Others have discussed smart cities as part of a larger conversation around the concept of “big data,” pointing out that data is not a neutral entity [5, 65], as well as the potential violation of civil liberties associated with increased surveillance and data integration [5, 26, 67]. In particular, some have noted the power imbalances that can arise in public–private partnerships, and that corporations are interested in large-scale data collection due to the ability to turn urban data into profit rather than to provide services related to the public good [32].

Urban sensor networks are not apolitical infrastructures—they are situated in particular physical places and social structures. Our work is in conversation with these critiques, providing both a grounded commentary on a particular smart city implementation and a framework that can help reveal how people influence and are involved in the flow of data through such systems. We aim to elaborate the simplified views of such systems in the engineering literature by discussing their sociotechnical nature. A sociotechnical perspective highlights the contextually embedded mutual constitutions of people and digital technologies—“the interdependent and inextricably linked relationships among the features of any technological object or system and the social norms, rules of use and participation by a broad range of human stakeholders” [54]. A sociotechnical perspective on urban sensor network deployment calls on us to examine how a multitude of human interests and organizational structures influence their construction. Carvalho's work demonstrates how this kind of sociotechnical perspective can provide useful insights about how technologies are tested and the ways in which social systems respond when smart cities are developed "from scratch" [7]; here, we use a similar approach to understand the impact of these same forces on the low-level sensor networks and associated data flows that underlie a particular city's ongoing smart city implementation and integration efforts.

2.3 Empirical and design investigations into smart cities

Our work is most closely aligned with those who have called for in-depth empirical work on existing, situated smart city implementations [26, 55, 64]. Shelton et al. point out how "smart city interventions are always the outcomes of, and awkwardly integrated into existing social and spatial constellations of urban governance and the built environment" [55]. Empirical investigations uncover how these "awkward integrations" occur in specific cities, and thus help disambiguate the smart city concept and provide a basis for grounded critique.

A similar investigation by McMillan et al. explores how data is used in governance in a number of Northern European cities; however, they focus on applications that involve citizens as data collectors or consumers, rather than urban sensor networks [33]. Others have looked at how databases are used across multiple scales of public sector work [30], examined the uptake of data-driven systems by police departments [62], investigated the bottom-up processes used in a local co-design projects in Taipei [21], assessed effects of the regional governance structure in the Boston metropolitan area on opportunities for creating a large-scale smart-city ecosystem [27], surveyed those implementing smart city projects across a number of cities [46], and critiqued the technocratic underpinnings of smart city planning and its capacity to support urban sustainability and ecological objectives [66]. These authors have engaged with the various motivations for developing data-driven infrastructure at a local scale and reflected upon both the power of data and the ways that data is influenced by power dynamics, including those between city residents and governing entities or between local and regional governments.

Authors have also tackled the relationship between data and place. Through an investigation of a number of data-driven projects on a single city street, Taylor et al. discuss how "small worlds" of data are created due to the boundaries of both physical and social geography [60]. Freeman et al. bring the concept of "placemaking" from geography to bear on the design of smart cities, highlighting the importance of understanding the iterative human processes of sense-making involved in creating places [21, 49].

Design and systems investigations in the civic space have largely focused on developing ICT solutions that engage residents with local policy making or advocacy, often via tools or workshops that allow residents to explore urban data sources [3, 14, 16, 20, 22, 24, 42, 51]. In addition to data consumers, city residents have also been treated as sensors or data sources through participatory sensing [1, 6, 28, 29] and hyper-local polling [11]. Design researchers have also developed ways to contextualize sensor data in order to make it more actionable [15, 50].

Our investigation provides an additional deep description of a smart city implementation in a particular geographical location. Our analytical approach complements and expands upon prior work by tracking how sociotechnical factors influence the end-to-end data flow in a number of urban sensor network systems. We construct a framework based on cross-cutting issues that is likely generalizable to other locations and could provide an appropriable lens for future empirical, design, or systems work in the civic space.

3 METHODS

3.1 Research context

We focused our investigation on projects in Denver, Colorado, a fast-growing city in the Mountain West region of the United States. The majority of projects discussed were explicitly part of the Denver Smart City program [9]. This program started in 2015 as part of an application to the U.S. Department of Transportation Smart City Challenge for mid-sized cities. Denver was one of seven finalist cities that received funding to develop a detailed smart city vision [61]. Although they did not ultimately win the competition, the initiative has since continued, primarily with projects focused on air quality and transportation.

Projects associated with the Denver Smart City program have been supported by various funding sources, including a Bloomberg Philanthropies grant for deploying air quality sensors in Denver Public Schools [10] and a federally-funded Advanced Transportation and Congestion Management Technologies Deployment Program (ATCMTD) grant for deploying pedestrian detection and vehicle-to-infrastructure communication technology [9]. The city also collaborated with Panasonic on the Peña Station NEXT development project, with the goal of constructing a smart community from the ground up [47] and is developing an Enterprise Data Management (EDM) system built on Microsoft Azure for managing Internet of Things (IoT) data across multiple projects [9, 35].

Other sensing projects discussed in interviews were not formally part of the Denver Smart City program, including projects to monitor water quality in city lakes and streams and the flood monitoring and response system operated by the Mile High Flood District. The Mile High Flood District was established in 1969 to coordinate flood mitigation and response in Denver and surrounding counties [36]. Participants also discussed projects supported by the Colorado Smart Cities Alliance, which facilitates regional public–private partnerships, including several projects in Denver [13].

3.2 Participants

We conducted interviews with 12 individuals involved in designing, implementing, and/or operating urban sensor network projects. We systematically recruited participants in order to provide perspectives across a number of overlapping sensing projects in the region. Most participants worked in (or had previously worked in) mid- to upper-level management positions within local government and therefore could talk about both the specifics and day-to-day operation of projects as well as the associated higher-level motivations, constraints, and organizational context. Half of the participants were city government employees working in the public health, public works, technology services, or transportation and infrastructure departments in Denver. Two participants no longer worked at the city—one had left to work in the private sector on smart transportation, and another had left to work at a regional coalition-building nonprofit in the smart city space. These participants spoke both about prior work with the city and their current work on projects relevant to the study. Two participants were employed by the Mile High Flood District, and two were academic researchers who had been involved in government-affiliated smart city projects.

3.3 Data Collection

We collected data via semi-structured interviews. We dynamically customized the interview protocol to prioritize the topics most relevant to each participant's expertise and experience. Each interview covered the following themes: the flow of data in specific projects, how those projects related to the broader concept of a smart city, and any other relevant experiences about the use of data in the participant's work. We prompted all participants to elaborate on both the technical as well as the social and organizational aspects of each project.

Interviews were mostly held at each participant's workplace, though out of convenience for participants, four interviews were held over the phone and one interview was held at a coffee shop. Interviews lasted 52 minutes on average. We transcribed all interviews for subsequent analysis, which was interleaved with ongoing data collection.

3.4 Analysis

We conducted a multi-phase analysis of the interview transcripts, starting with deductive analysis using an initial set of codes relating to the process of moving from data collection to action. Our coding scheme included four tags: *data collection*, *data management*, *data analysis*, and *data impacts*. We adapted these tags from Naphade and colleagues' foundational work discussing the "system of systems" vision of a smart city and related challenges [38]. In much of the engineering literature related to smart cities (see also e.g., [8, 12, 56]), these four phases are envisioned as part of connected, closed-loop systems which are automated by technical infrastructure and then connected into an overall "system of systems" via shared databases or analyses (see Figure 1).

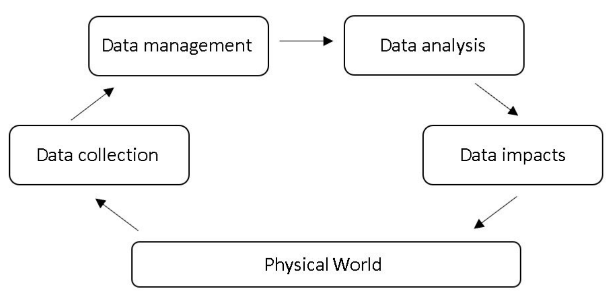

Fig. 1. Closed-loop coding scheme adapted from Naphade et al. [38]

Initial coding scheme:

- Data collection: *collecting data from sensors or people's personal devices*
- Data management: *storing data, primarily in databases that are shared across multiple departments or agencies*
- Data analysis: *processing data into actionable information via both real-time and long-term analysis, as well as developing prediction models*
- Data impacts: *influencing outcomes via informing, actuating, responding, or controlling*

In the first round of coding, the first author applied one of these initial "Process" codes and a simultaneous "Project" code to each relevant excerpt in the transcripts [37]. The "Project" code allowed us to keep track of projects that were discussed by multiple participants. Applying these two codes to each relevant excerpt allowed us to look both within "Process" tags, themselves, as well as at the flow between those tags across specific projects. During this first round of coding, we also wrote analytical memos [53]. In these memos, we reflected upon how each of these initial codes fit to data. In some cases, we chose to expand or elaborate upon definitions of the initial tags based on our sociotechnical—rather than primarily technical—perspective.

In the second round of coding, the first author then performed multiple rounds of iterative, inductive coding within the data for each of the "Process" codes to identify sub-codes within each phase of the smart city process. Given the exploratory nature of this study and desire to understand the flow of data, codes were primarily descriptive in nature (i.e., open codes [59]), for example, identifying when a participant discussed sensor placement or developing a custom data platform [53]. In this round of coding, we also added additional high-level process codes for when participants discussed "project motivation" and "project funding," which were not part of the initial four codes but emerged as significant recurring themes. A final round of axial coding was performed to group similar sub-codes.

4 RESULTS

4.1 Data collection

Participants brought up many technical aspects of developing and maintaining data collection infrastructure. For example, many participants discussed the relative quality of different sensing technologies and setting up protocols to ensure that a sensor network was robust to failure. However, we found evidence of a broader sociotechincal narrative even in simple choices such as what sensors to use and where to place them. Such choices are made within an organizational context and decisions needed to align with existing budgets, collaborative structures, and spheres of ownership.

Choosing sensing technologies typically involved weighing trade-offs between sensor cost and data quality, with participants needing to make judgements early on about the level of sensing accuracy required for particular use cases. In some cases, sensor cost was a prohibitive factor, leading participants to continue manual data collection instead of investing in automated systems. Our participants discussed placing sensors strategically in anticipation of future analyses based on existing organizational collaborations. For example, in an air-quality sensor network, sensors were deployed at public schools where it was possible to also collect asthma-related health data, and near street intersections at which existing sensors monitored traffic flow.

In some projects, such as flood and air-quality monitoring, the goal was to deploy sensors in a way that provided distributed sensing capabilities across a geographical area, often extending beyond the boundaries of a specific city. In other cases, sensors were placed in specific locations based on the requirements imposed by funding sources or to monitor specific construction projects, traffic intersections, bodies of water, or pilot projects.

We came to see the deployment of sensor networks as just one of the various, often-relational activities undertaken by those working in local governments to obtain data, and considered by our participants to be part of the smart city concept. We asked participants if they obtained data from other entities as part of their work, and participants often discussed such collaborative data collection activities unprompted. Data was shared between departments or other governmental entities via shared repositories or interpersonal connections:

> *So we can get it through the data-sharing network. We also work closely with a lot of those folks. So we can call them up and say, 'Hey, can I get this data?' and the data sharing network is difficult, so they will frequently say, 'Sure, let me pull it and send it to you.' So it's not as easy as going out to the cloud and just pulling the data down. But you can...it's accessible. You can get it.* (P3)

Participants also discussed getting sensor data from private entities via mandated reporting. For example, the City of Denver collected data from ride-sharing and scooter companies in order to understand the impact of these services on traffic patterns. Such data collection agreements demonstrate how a rapidly changing consumer technology ecosystem offers both challenges for local governments, in terms of being able to fully understand urban systems, and opportunities for collecting new forms of data. In addition to mandated reporting from some companies, sensing strategies capitalizing on widespread smartphone use are also in operation in Denver—for example, tracking traffic flow via Bluetooth sensors. Additionally, participants mentioned how specialized sensors, such as those for air-quality monitoring, are increasingly purchased by consumers, creating a bottom-up data collection infrastructure that city governments can also leverage.

Most participants saw automating data collection with sensors as the ideal, with the goal of continuously collecting and processing data such that the elapsed time between data collection and use would be minimized. However many participants described current practices that combined a variety of data collection strategies with different levels of automation:

> *And so we get, gosh, probably like 15 different sources of traffic data. So some of it is we go out and collect it...some of it is we have various cameras or sensors around the city.* (P8)

Ad-hoc infrastructure development and heterogeneous public and private ownership of assets also led to fragmented data collection. In the most extreme cases, the same data might be collected by multiple entities for different reasons:

> *And then the state, the Department of Water Resources or [Department of] Natural Resources...they have hydrographers who go out and survey channels and then collect the same data that we're collecting.... But they're doing it to determine water rights compliance issues.* (P4)

Our broader sociotechnical lens also allowed participants to discuss what data might not be being collected and why. For example, Participant 5 discussed the possible hesitancy around putting air quality monitors at schools, given that schools can be placed in highly polluted areas: "*I mean some cities don't want it also because what they have seen with traffic patterns, we have also been horrible at placing schools because it's cheaper land like right in there*" (P5).

Examining the choice of sensor type and placement from a sociotechnical perspective revealed the multiple constraints under which cities operate when developing data collection strategies, including those imposed by budgets, city borders, existing infrastructure, data protections, and private ownership of data collection assets. Such constraints can lead to data collection infrastructures that offer partial or duplicated coverage of phenomena of interest and that are heterogeneous in their levels of automation and their ownership across public, private, and academic sectors.

4.2 Data management

The engineering literature around data management in smart cities has idealized centralization projects, in which data sources are integrated across multiple domains [38]. Our participants often discussed such centralized, cloud-based data management solutions as an ideal, and mentioned challenges with such projects that have been identified in the engineering literature, such as in developing unified data models across heterogeneous data sources and types [51].

In practice, centralized database projects were pursued at various geographic scales and levels of cross-domain data integration. Participants discussed data management—the databases and associated schemas used in projects—in the context of supporting organizational structures, collaborations, and processes.

All participants working on smart city projects in Denver government mentioned the Enterprise Data Management (EDM) system, which was the city's attempt to centralize data storage from a number of departments and sensing projects into a single data platform. This project, completed via a partnership with Microsoft [35], was a way to "*de-silo*" (P1) datasets across departments and was "*the heart of*" (P1) the cross-collaboration across departments in Denver's Smart City program. As the EDM project rolled out, it initially focused on storing data from specific Denver Smart City projects, as building the EDM from scratch took a lot of human effort and technical expertise:

> *Right now, what I have to work with is a lot of disparate systems... so we have to go from one system to the next system, the next system. So with part of this, with the new datasets that we're bringing in, I don't want to add one more system to our systems. I want an*

> *overall system that's gonna manage those. So we are building our own Azure instance of an Enterprise Data Management system.* (P10)

Participants also discussed custom data platforms that were used to centralize data within a specific domain, such as transportation, air quality, or flood monitoring. In one of the cases discussed, an air quality sensor network sent data to a custom platform before it was also sent to the EDM. Similarly, traffic data used to be stored in "homebrew databases" [63]—"*in various Excel sheets or PDFs*" (P8). While those working with traffic data intended to eventually move all data into the EDM, it was higher priority to move data into a custom, department-level platform that was specialized for traffic data management:

> *So we had the option of putting a lot of traffic data into the EDM...and we went with a cloud-based solution that's already commercially available that we could buy. The big thing with the EDM is that you kind of have to figure it out all yourself...you know, how do you access it? You gotta build a website around it and what does it do here and what does it do there? So you're literally kind of starting from scratch.* (P8)

The Mile High Flood District had their own custom data platform which integrated multiple relevant data sources at a regional scale around the common purpose of storm monitoring. Participants working on this system were resistant to the prospect of integrating data into or from this special-purpose data platform if it was not essential to the Flood District mission.

Data was also accessed from, stored in, or sent to regional, national and international databases. Data management across large geographic scales or heterogeneous sensor ownership was made possible through the use of data formatting standards, which were developed and enforced by organizational units. For example, the Mobility Data Specification [44], promoted by the Open Mobility Foundation (OMF) [45] simplified obtaining data from ride-sharing, scooter, or other mobility companies. The OMF is run by a coalition of municipal governments and, among other things, develops and promotes standard formats that private companies use to report mobility data to cities:

> *Last I looked, there were probably a dozen cities that are part of OMF and so they're basically saying like, we are the public sector and we should govern how the public right-of-way's used and the data is collected so we can actually use it in a standardized format. Whether you're in LA or Pittsburgh. And yeah, they also have private sector members, so the private sector sort of pays a fee to be part of the OMF. And the OMF is governed by these public sector entities.* (P11)

As in data collection, participants discussed the complications of negotiating private vs. public ownership of and access to databases. Participant 1 described the challenge of finding vendor partners who were willing to cede ownership of stored data:

> *Because we as a city, we want to have ownership over the data. And vendors, obviously data is powerful...so like then they want to keep ownership of that data...and I think it's something we as a smart cities group need to focus on in terms of data control type of policy. Just to ensure that all the data we are getting, like we have ownership over it and it doesn't reside in someone else's database...because at any time they just could just cut off access and then we're just [out of luck].* (P1)

Participants also mentioned privacy considerations around storing data, which was particularly an issue for video footage. For example, video footage in the traffic management center was accessible in real-time but not stored.

Discussing data management with participants revealed the importance of organizational structures in the creation and maintenance of databases. Data standards, which were created and enforced by institutions, simplified integration of data from sensors that were diverse in ownership and locations, particularly within and across jurisdictional boundaries. Cross-domain data integration within databases was identified as technically challenging and time-consuming, but in some cases it was valued for the collaborations it encouraged. In others, special-purpose databases were preferred over databases that stored data from various projects or domains.

4.3 Data analysis

Participants discussed data processing and analysis activities that occurred at various time scales and with different levels of human effort or expertise. Before undertaking analyses, moving from a raw, sampled medium to a dataset could involve human effort and decision-making. In an extreme case, water samples needed to be transported to a laboratory and cultured in order to obtain an E. coli level—a process that eliminated the ability to perform real-time detection of bacteria levels.

Similarly, in the case of video detection, a person or algorithm obtained data of interest, such as the number of cars passing through an intersection, from raw video footage. In some cases, variables of interest were estimated based on the value of a more easily-sampled variable. And, as a general concept, sensor measurements were used to estimate what was happening over a broader geographic area—a technique that could miss important spatial variation if the level of spatial aggregation was chosen inappropriately. In some cases, private entities provided data to which proprietary data processing techniques had already been applied.

Participants also discussed various methods to detect erroneous data. Just as human subjectivity was involved in data clean-up and processing, the ability to detect faulty data could depend on deep familiarity with the data, such as knowledge of common indicators of sensor failure, or knowledge of what other datasets could be used to corroborate sensor readings: "*If you're used to dealing with the data and you understand it in depth, you probably have a real sense for what's probably wrong. The average person wouldn't know*" (P9).

In the projects we discussed, automated processes built into custom data platforms were used primarily to perform simple calculations or detect when a variable exceeded a threshold. However, participants repeatedly discussed higher-level analyses being performed by people who were experts rather than through automated processes. In many cases, these experts were external to the government agencies directly involved with operating and maintaining the smart city infrastructure. For example, academic researchers collaborated on projects that produced data of scientific interest:

> *So finding and working with other people who are interested in what that data can do. And basically do data analysis on it...cause, I mean, I'm one person.... So I mean just kind of like we're basically creating this like research infrastructure for anyone to come in and take all the data and run with it.* (P1)

In some cases, experts either external or internal to the government were "built into" systems as permanent fixtures rather than as temporary collaborators, as discussed by Participant 4 in the context of the flood monitoring network: "*...[Data are] also transmitted to two consultants that we use to kind of do some data processing...so we try to split up the load a little bit to give different consultants some work*" (P4).

In many cases, decision-making was facilitated by modelling and prediction based on historical datasets rather than analysis of contemporaneous data. Participant 3 discussed how the dominant focus of other smart city projects on real-time data analysis was not in line with the majority of analysis work required for water quality monitoring:

> *So they want to have all these different real-time data sources out there that they can start to look at them and see, you know, this is happening and it's causing, you know, the side effect is this. But what I need to be able to do is I need to be able to kind of look backwards*

> *to understand what's been going on so that I can think about going forwards. How do we adapt to it? ... You know, we have close to 20 years worth of data and we're actually just now getting to the point where we can really start to look at it and understand what's going well.* (P3)

In addition to looking at long-term trends, participants discussed a general pattern of collecting baseline data, implementing changes, and then evaluating the effect of those changes. This pattern complemented a broader move towards data-driven policy making, in which the city attempted to make changes based on data analysis rather than assuming best practices would work.

Generally, the discussion of data processing and analysis using a sociotechnical lens exposed how human subjectivity and effort was involved in nearly all stages of moving from a raw sensor reading to an output of interest. This discussion also exposed how city data infrastructures must support multiple analyses that operate a different temporal scales including for real-time alerting, modelling to enable prediction of future events, and analyzing "before" and "after" data to evaluate the effectiveness of interventions.

4.4 Data impacts

Participants mentioned many ways that information produced from sensor data impacted the world, or ways that they hoped or predicted it would. While some of these impacts happened quite quickly—for example, in response to crises—others happened over long periods of time through the accumulation of new local knowledge. While some impacts were realized via automatic or remotely controlled actuation of infrastructure, the vast majority were realized by informing either government decision makers or members of the general public. Participants discussed many ways in which they anticipated people would respond to this information, including infrastructure maintenance, planning, or law enforcement activities by government actors; and personal health or safety-seeking, environmental stewardship, or advocacy activities by members of the public.

The most immediate, automated data impacts were achieved via actuation of infrastructure due to patterns detected in data streams. For example, detecting when people were in a crosswalk and using that information to adjust traffic signal timing. However, as mentioned by Participant 9, it could be risky to rely on a fully automated response system: "*but I'm going to want some human intervention there not just gauges thinking they know the answer*" (P9).

Other nearly immediate impacts on the world were achieved once people internal or external to the government were alerted to patterns or thresholds in the data that required response. For example, in the flood monitoring network, an alert was sent out to response coordinators if precipitation levels exceeded a threshold. Responses to alerts also varied in their level of human effort—some could be achieved via remote control of infrastructure, thus emulating what might be possible with an automatic system with additional oversight. More complex responses might involve sending people out to respond to an event—for example, people might be sent out to close a flooded bridge or on snow plows to clear snow.

Sensor data also contributed towards a regularly updating body of knowledge that was used to make longer-arc decisions about city operation, often accessible via online dashboards. For example, ride-sharing data or average travel time data might be used in long-term traffic planning. Alternately, air quality data could impact decisions around permitting or building public projects. New information could be distributed via other artifacts as well, such as summary documents, scientific publications, or conference presentations.

Across the projects we discussed with participants, the desire to use data to inform the public was a repeated theme. Flood-related alerts go to members of the public as well as first responders. Online dashboards used by government planners were often also fully or partially accessible to the general public. In the case of the air quality sensing project, in-situ dashboards displayed data to those at schools with air quality sensors.

In discussing desired impacts, participants often assumed that the public would both access data and use this information to alter their behavior. For example, traffic issues might be mitigated if people planned trips based on traffic data. Or, people might decide not to swim in a stream if they saw that E. coli levels were high. Participants also discussed how seeing some environmental data might inspire people to be become better environmental stewards or get involved in political advocacy. Participants also anticipated possible unintended consequences of relying on personal decision-making. For example, people might intentionally put their property in harm's way if it provided them a financial benefit:

> *Now there's people that want to get in hailstorms so they get their car wrecked so they can get a new paint job. Let's say they've been dinged up a lot. They just want to collect the claim. So I could see it being used in unintended ways.* (P9)

Participants acknowledged that simply providing access to data did not necessarily lead to data impacts, as the public's success with sense-making around this data and using it to alter behavior in the desired way was not guaranteed. In some cases, this inspired the creation of supporting educational programs that encouraged public engagement around sensor data. Our discussions revealed that while some data impacts were immediate and automated, many relied on human interpretation, complex decision-making, and collective action and thus were difficult to fully anticipate or structure.

4.5 Project-oriented view

An additional thread emerged during the interviews that conceptualized the infrastructures discussed as *projects* that needed to be planned, funded, gradually implemented, and iteratively evaluated. Participants repeatedly discussed the goals of smart city projects and how they intersected with funding opportunities. They also discussed ways that infrastructures were designed and implemented such that they could be continuously evaluated against these goals. This project-oriented view reinforced the idea of projects as boundary objects [58] around which planning, funding, design, deployment, maintenance, and evaluation activities could be carried out across collaborators and constituencies.

Participants discussed various ways that data could be collected to evaluate the impacts of smart city projects, for example by asking those consuming information from such infrastructures to fill out surveys about how they used the data, or by setting up ways to relay telemetry data from sensors back to storage. Such data was collected so that participants could prove that new technologies were useful. "*But we're also trying to evaluate, 'can we really do this in a way that's useful?' So those are the kind of questions that when we're trying out something new, we test it and we start asking those hard questions, 'Is this useful? Is this useful?'*" (P9).

Many smart city funding sources mandate reporting about project outcomes and technology use. Evaluating sensor networks incrementally was common, supported by the tendency to deploy sensors gradually in strategic locations. In other cases, the goal of smart technology deployment was never to create an entire, functional infrastructure, but rather to deploy pilot projects in order to understand whether a particular technology could feasibly be used in the future. Participant 11, however, critiqued the tendency to focus on pilot projects and isolated deployments of smart city infrastructure rather than approaching infrastructure development from a holistic perspective:

> *I think the biggest challenge that we're facing overall is asking this question is what's most important from an outcome perspective that we're trying to accomplish and then figure out like, okay, if that's the outcome you want, then what are the... where's the problem?*

> *Or something very specific that you can solve. I think what you’re seeing is a lot of pilot projects, things that are very tactical, but we’re not thinking far enough through.* (P11)

Some aspects of smart city projects frustrated efforts towards traditional project evaluation. One challenge came from the scope and diversity of actors within a single project. Cities have limited leverage and are involved in projects largely due to their jurisdiction over public streets, land use, and zoning within their geographic boundaries. Due to the end-to-end complexity of sensor networks, smart city projects often involved a constellation of actors spanning sectors who might come together due to mutual interest, rather than around a government-defined goal:

> *Yeah, so that’s another complexity of smart city solutions is that if you look at everything it takes to put one together, you have a lot of parties that need to be involved. So just as an example, this project... So you have, starting at the end...you have sensors that measure water quality.... Cisco provides the network and the gateways that those sensors will talk on. And then there needs to be some sort of platform that makes sense of all the data that’s coming from the sensors through the network to a user. And generally that is a different company developing the dashboard user interface and the IoT integration of all the data to make sense of it.... So there’s potentially four companies involved here—at least.* (P12)

Traditional government processes of collecting community feedback were also more challenging in the case of smart city projects, due to the invisibility of infrastructure [57]. Examining this thread in the transcripts revealed the challenges involved with evaluating smart city projects and how the gradual roll-out and evaluation of such projects happened simultaneously with their operation.

4.6 Smart cities

As part of our interviews, we asked participants to reflect on the specific term “smart city” and how it did or did not relate to their work. Most participants considered their work to be related to this term and did not protest our focus on urban sensor networks as opposed to other urban ICTs. However, participants most frequently talked about smart cities through the lens of *collaboration*, rather than through the lens of specific technologies. Some discussed the Denver Smart City initiative as removing “*silos*” (P1, P8) between departments. Others focused on public–private partnerships with technology vendors. Participants also mentioned learning from other cities, participating in smart city industry events, and the importance of regional collaborations.

Some participants expressed frustration with the smart city term, characterizing it as ambiguous, a “*buzzword*” (P11), or part of an industry “*hype machine*” (P10). Multiple participants discussed the term as a form of branding which had come into vogue in response to funding opportunities. Despite this frustration, participants saw promise in enabling new approaches to data-driven problem solving. Even those who were frustrated with the ambiguity or trendiness of the term acknowledged the opportunity to use new technologies to gather “*better data*” (P5).

5 DISCUSSION

A sociotechnical perspective helped us discuss with participants how human decision-making and actions influence and enabled the flow of data within urban sensor networks, as well as the ways that organizational structures constrained and were built in parallel to these technical structures. In our emergent analytical framework extending Naphade et al. [38], urban sensor networks can be broken down into the phases of data *collection*, *analysis*, *management*, and *impacts*. Each of these phases can display variability in *immediacy*, *automation*, *geographic scope*, and *ownership*. Differences in how these constructs are realized across phases or projects can often times give rise to sociotechnical frictions—or even lead to fundamental technical incompatibilities across systems.

Each urban sensor network can also be conceptualized as a *project*, which is a boundary object [58] facilitating collaborations between various organizational entities and people.

This lens helps us form a grounded critique of the smart city concept. In particular, our work suggests that one-size-fits-all urban sensor network solutions will need to confront the particularities of each city's geographic features and constellation of public, private, and academic stakeholders. Inflexible, generalized solutions will face challenges and may fail due to each city's unique nature. Additionally, based on our participants' perspectives, constructing a fully-integrated "system of systems" may not always be the best investment for solving problems in cities. Any large-scale projects that aim to deploy sensors at scale or merge various sensor networks will need to respond to the specific sociotechnical context surrounding their implementation and evaluate various cost–benefit trade offs in light of local policy priorities. They will also need to consider potential public concern around data surveillance [20, 32, 67] and whether the motivations of public, private, and academic entities involved are aligned in such a way that will allow the project to meet its goals. In some cases, smaller-scale, targeted sensor deployments, or even decisions to rely on less immediate or less automated solutions, may be preferable to archetypal "smart city" efforts—not only for those living in the cities, but also for those civil servants and contractors charged with the design, implementation, and maintenance of its infrastructure.

In the following sections, we discuss how each of the cross-cutting constructs in our analytical framework emerges via a sociotechnical lens and contributes to this grounded critique.

5.1 Immediacy

The concept of "real-time" is core to smart city rhetoric [26], and immediacy can be explored in the duration of and between phases in the smart city process. In systems that are closer to "real-time," data is captured from sensors nearly instantaneously at regular, short intervals of time, automatically uploaded to databases, and integrated with other data streams. Impacts, too, can be extremely immediate—for example, when data thresholds cause automatic actuation or informational alerts that call for immediate human action. In some cases, technical constraints or the need for human decision-making and interpretation cause stages to be *less* immediate—for example, collecting E. coli data requires a sample to be processed in a lab, thereby increasing the time between sample collection, analysis, and impact.

Immediacy is also influenced by the time scales of phenomena of interest. Flood warning systems are oriented around weather events and must operate extremely quickly, whereas data-driven systems supporting transportation planning or water quality analysis operate over longer time scales. In the latter two cases, participants were interested in responding to phenomena such as population growth, which happen over the course of years or decades. Impacts, too, can vary in immediacy, with changes to policy or development of city infrastructure taking a long time and requiring the input of many stakeholders. In many cases, city governments partnered with academic researchers, for example, to better understand the relationship between air quality and health. Scientific findings are produced over the course of months or years, and may have little to no immediate impact on policy.

Longer-duration analyses and resultant non-immediate impacts were common across the projects that we discussed with participants, even in cases where systems did support near–real-time access to data. This leads us to think about how the longer arc of such systems might interact with other long-term processes such as natural and social change, political terms, and grant opportunities, as well as the gradual roll-out of smart city projects, themselves. Real-time systems are enabled by technologies such as sensors, fixed and wireless networks, and algorithms for analyzing large datasets on the fly [23]. However, immediacy may not be an essential element of a data system used to inform policy reacting to longer-term processes such as climate change or urbanization [55].

Further, mis-matches in immediacy between sensor network implementations and policy-making lead to the kinds of questions raised previously by P3 when critiquing the role of real-time data in the traditionally retrospective domain of water quality monitoring: *"I need to be able to kind of look backwards to understand what's going on...."*

5.2 Automation

Related to the concept of "real-time" is that of automation, or whether human activities in each phase have been replaced with those of machines. The extent to which automated actions or decisions require human oversight has been core in critiques and re-imaginings of smart cities [20]. While automated data collection can and does utilize installed sensors continuously sending data, we also talked with participants about instances in which humans manually collected data—for example, performing *in situ* traffic counts. In other cases, human effort was required to process raw data—for example, interns performing traffic counts manually (and, at times, asynchronously) by watching video streams. This activity, of counting cars and pedestrians, is in the process of being automated via computer vision software, and many instances of human effort in closed-loop systems were seen as temporary workarounds while participants waited for funding and technology to become available for automating these processes.

However, we also found that human oversight was necessary across all phases, even in mostly automated systems. Familiarity with particular sensors (e.g., knowing the quality/consistency of the data provided by the sensor, understanding the degree to which the data provided by the sensor was pre-processed or aggregated, for example, when passing through the sensor provider's commercial data platform) and hyper-local geography ("at intersection x," "near the stadium," etc.) was needed for error detection, data analysis, and for determining appropriate interventions. Human expertise was required for performing complex analyses, such as combining data across natural and social subsystems, and for interpreting the output of prediction models, such as those anticipating the effects of weather events. Complex analyses and interventions also required supporting organizational structures—ongoing inter-departmental working groups or regular calls with academic experts. Further, the dominant mode of achieving "data impacts" via information dissemination and its secondary effects relied on human sensemaking over data. This final process of sense-making in and of itself is a complex and rich area for design [50].

Finally, the gradual roll-out and evaluation of sensing systems required human decision-making throughout—in choosing industry partners and pursuing funding sources, in strategically placing sensors, in developing data models, in choosing if and when to use which data standards, in choosing alerting thresholds, and in designing data dashboards. Trust in automation was built up gradually over time, as systems rolled out and became operational. Given the safety-critical nature of many applications, human-in-the-loop configurations, in which humans provided some oversight over automated processes, were common (see [31]) and may be necessary.

5.3 Geography

Different phases within a project may also operate at different geographic scales. Data collection may be constrained to city boundaries, or it may be done regionally or nationally. Databases, and associated analyses or interventions, operate at different or across multiple "scales of influence and accountability," as discussed in a human services context by Le Dantec et al. [30]. Broadly, natural boundaries and political boundaries do not align—many participants discussed how you need a regional view to understand flooding risks, air quality, or water quality within a city.

The importance of "social–ecological alignment," in which human collaborative structures mirror ecological connectivity, has been stressed within the environmental studies literature [4]. Environmental sensing projects may have greater success if they span a large enough geographic area to support this alignment, or if they are easily combined between neighboring territorial units that are connected ecologically. Cities have jurisdictional responsibilities over streets and zoning laws, suggesting that coalitions need to be built among neighboring cities. In our research context, this was observed in the Colorado Smart City Alliance [13]. Creating these necessary regional partnerships in conjunction with technology deployment increases the organizational challenges associated with developing and rolling out smart city infrastructure.

Natural and political geography also influence the extent to which smart city projects must be individualized for specific cities. Our participants brought up multiple instances in which local quirks such as a dry climate, tendency for winter hail storms, or dam controls decreased the efficacy of smart city technology solutions that worked elsewhere. Lack of standardization in the realm of urban IoT data models [48] also means that many cities have developed their own *ad hoc* technical infrastructures. Those working in cities and interested in innovation spend time trying to learn what has worked in other places, but must constantly try to adapt such solutions for their local natural environment, legacy technology systems, and city policy priorities. This introduces tensions in public–private partnerships, as private companies are interested in solutions that can be re-deployed in multiple cities rather than in creating custom solutions so as to leverage economies of scale [26]. This tension also leads to cities going through the process of deploying new technologies simply to try them out, confusing motivations between trying to create an operational system and evaluating new technologies via pilot projects.

5.4 Ownership

We also observed that varied institutional ownership occurred within and across phases in the smart city process. Within projects, there might be multiple public, private, and/or academic entities involved in various aspects of data collection, storage, analysis, and impacts. Public ownership often breaks down along departmental lines within cities, and then at different geographic scales. Different industry partners are brought in for different parts of an end-to-end system based on their specialization (sensing, communication infrastructure, cloud storage, etc.). Academic partnerships are typically pursued in order to apply for research grants to support technology development that underfunded cities wouldn't otherwise be able to support financially, and multiple academic partnerships might be associated with the same sensor network. Finding common ground across multiple entities with different core motivations (public service, profit and R&D, scientific research) can at times lead to power asymmetries [32] or sub-optimal configurations—for example, the city committing to develop transportation infrastructure at a smart development despite its low initial population density.

Private entities also create sensing systems without actually conferring with the city, leading city governments to adopt reactive strategies—for example, mandating data reporting from scooter companies. Tensions can also arise when private companies want to maintain ownership over data. Cities don't have the funding and human resources to build and run entire end-to-end systems, but are also concerned that they might lose oversight over these systems without suitable data ownership agreements. Private companies have a huge financial incentive to maintain ownership over data from urban sensor networks [52, 67]. Ownership over data, especially that coming from video streams, can also introduce questions around privacy and secondary uses of data, with our participants noting potential public concern over both public or private ownership of such data.

Different domains (air quality, water quality, transportation, etc.) are understood by different constellations of departments, regional coalitions, experts, and private companies. Thus, when centralization is attempted across sensing projects (in the "system of systems" vein [23, 38]), the overall number of actors involved increases. Large projects supporting multiple purposes also run into the need for evaluation across different metrics; many of our participants pointed out the necessity and challenge of articulating, from the city's perspective, why a particular smart project is pursued as it morphs over time due to the inclusion of different funding sources and partners. The desire to evaluate or justify smart systems leads to additional work to build in telemetry and reporting systems. This finding also highlights how policy goals related to public well-being of a smart city project can become muddled over time.

5.5 Tension between cross-domain integration and fragmentation

The phrase "system of systems" captures the idealization of cross-domain data integration: that combining data from different domains (i.e., air quality and transportation) can lead to insights into the relations among sub-systems, which can improve models of the city and promote data-driven policy making. [23, 26, 38].

In the city that we studied, we did not find an entirely integrated "system of systems" [23, 38], nor did we find many issue-oriented "small worlds" of data, to borrow from Taylor et al. [60], nor domain-fragmented "big little data" projects, to borrow from McMillan et al. [33]. Our sociotechnical investigation allowed us to unpack various factors pulling both towards cross-domain data integration and in the opposite direction, towards fragmentation. We discussed this tension briefly in reviewing our participants' discussion of data management. Although data management is core to this discussion, it is not the only potentially integrated or fragmented phase in the smart city process, and issues related to cross-domain data management arise due to sociotechnical factors operating in other phases.

Cross-domain fragmentation has been investigated in the urban space from a technical perspective—heterogeneity in IoT architectures makes it challenging to integrate data from different data systems [48]. Fragmentation has also been described within urban studies in the sense of segmentation and specialization of public space due to market forces, which can be reflected in targeted smart city projects [34]. City governments are themselves fragmented into domain-focused "silos," and regional projects span multiple jurisdictional boundaries, each with their own unique "silo-ed" structure [27]. It is not surprising, then, that fragmentation by domain occurs in technical and organizational structures associated with smart city projects.

However, our investigation also uncovered other factors pushing towards cross-domain fragmentation, including the gradual process of rolling out and evaluating new technology, supported by opportunistic funding and public–private partnerships. Piloting new technology is an essential step in order to understand if it can operate in a city's climate, but this can also lead to a number of small, pilot deployments, in contrast to top-down deployment of an integrated system. Projects require large teams with different companies or external experts providing input within different domains. Data integration projects, such as the EDM, require significant human effort and organization to build from scratch.

These findings suggest that if cross-domain data integration is a goal, a multi-pronged approach is necessary—and that some areas may be more tractable for city governments than others. Cities can capitalize upon forces pushing towards data integration. Smart city rhetoric itself, along with associated funding sources and industry trends, can be strategically adopted by cities that want to organize in a way that supports cross-domain data systems. There is also interest from the scientific community in understanding the relationship between natural and social subsystems of urban environments, opening the doors for mutually beneficial partnerships between cities and university researchers. Cities can also become involved in the collaborative development of technical standards that reduce the human effort required to set up integrated data systems.

While participants were excited about combining data across domains, it is also worth reflecting on when data integration is ideal. Sensing projects with a clearly defined scope and set of success metrics might be more resilient and less unwieldy than general-purpose infrastructures. They might then be more tractably evaluated against public policy goals or better resist drift as new partners are involved. Our work also raises the question of whether cross-domain data integration affords sufficient public benefit to warrant the significant logistical and financial effort involved and potential risks such as public concerns over surveillance. Of course, there is no ideal strategy, and one of our findings is that there is no one-size-fits-all approach to smart city projects that will work across different political and geographic environments. Instead, the optimal degree of cross-domain data integration should be considered within a specific city context and in light of particular policy goals.

6 CONCLUSION AND IMPLICATIONS

Cities around the world are using sensing technologies to address challenges and try to improve quality of life for residents. We interviewed those working on urban sensor networks to understand how sociotechnical factors influence the flow of data in such systems. We used participants' words to flesh out a realistic description of how a smart city vision is partially realized in a specific city. By contrasting these professionals' descriptions with canonical understandings of a smart city from the research literature, we suggest generalizable sociotechnical tensions, challenges, and opportunities in the urban sensing space. While this work aims to expand on our collective understanding of smart cities, we also hope that the pragmatic examples offered provide some clarity within a space plagued by ambiguities, idealizations, and dystopian visions. In particular, our participants' fundamental understanding of their smart cities-related work as being about *collaboration* draws attention to this domain as being a rich site of study for CSCW researchers.

6.1 Implications for CSCW research

Our work suggests that urban sensor networks can and should be thought of as sociotechnical systems, and that empirical work investigating human–data interactions should not be confined to public-facing interfaces. Rather, we call for further investigation into how humans and organizations are involved throughout entire end-to-end urban sensor networks. This research highlights the importance of considering alternate units of analysis in the study of smart cities than have been used in previous inquiries—for example, demonstrating the utility of exploring the experiences of those individuals responsible for designing, implementing, and maintaining smart city infrastructure, rather than focusing more exclusively on study of urban policymakers, community members, or the resulting configurations of implicated technological artifacts and data networks. Such investigations can provide more concrete explanations for how data that is produced by such systems is influenced by human decision-making through time and across political, organizational, and geographic boundaries. Our emergent framework provides a lens complementing the existing, technically focused ways of examining smart city infrastructure that can be appropriated to form investigative protocols for future work.

Further understanding of these mechanisms is essential, as such systems continue to be taken up in safety-critical areas, and due to concerns over the introduction of human biases into big data systems [5, 43]. Our research is in conversation with work showing that data-intensive commercial systems can inadvertently reproduce societal biases (e.g., racial and gender inequalities [41]). Although our data is not appropriate for diagnosing whether such biases exist in the information produced from sensor networks in our study city, we provide a sociotechnical framework for future investigation into the flow of data in urban sensor networks. Future investigations might be able to leverage this framework to uncover, correct if apparent, or prevent biases in city policy that might be introduced or exacerbated by the uptake of city-scale sensing systems.

Design-focused research in CSCW (and human–computer interaction, more broadly) has largely been concerned with connecting the public to governmental decision-making processes and improving the process of human sensemaking over civic data (e.g., [3, 14–16, 22, 24, 42, 50]). Our work suggests that human–data interactions should be explored and supported in other parts of systems as well—for example, in decision support for placing sensors, in the design of city-scale sensor data models, and in the design of effective strategies for human oversight over smart city automation.

6.2 Implications for the design of smart city data platforms

Designers focused on human–data interaction probably won't be directly involved in the creation of new sensors or communication protocols. But there is increasing practical need for design of both domain-focused and general-purpose data platforms for smart cities. Based on our framework, designers should carefully consider the level of *immediacy*, degree of *automation*, *geographic scope*, and institutional *ownership* across the phases of urban sensing systems that feed into the analytic decision-support dashboards and alerting systems that constitute the "front-end" of smart city platforms. Designers should not consider the steps before data enters front-end systems as a mysterious black box or the data itself as essentially objective. Rather, designers should think about how to reveal potential sources of error or subjectivity involved in the sensing, aggregation, and processing of data [18].

Similarly, designers should consider the context in which data is consumed and the subjectivity involved in interpreting (often visually-presented) data, providing explicit support to help people translate data into action (bad air quality means make x decision about recess, for example), or to perform long-term or before/after analyses (e.g., *What happened last year at this time? What happened before and after* x*?*). Additionally, designers should consider the broader context of ownership in which a data platform is embedded and the various motivations of actors involved in producing and using the data from such a platform.

6.3 Implications for city policy

Our work also revealed and suggested some municipal-level policy strategies that can act to decrease ambiguity and improve outcomes associated with smart city projects. First, it is essential to decide what the goals of such projects are. Is the goal to create an operational system that improves health or safety for city residents? Or is the goal to test some new technology? Even if projects are pursued opportunistically based on funding sources or available partnerships, having a clear articulation of goals and what success looks like decreases ambiguity associated with a project.

Other useful strategies include pursuing regional collaborations, as well as following or collaboratively developing data standards. Cities can also develop clearer policy around data ownership and protections and hire staff in a way that increases their technical capacity. Broadly, cities should consider when pursuing cross-domain data integration projects provides value, and in what cases fighting the forces pushing towards domain fragmentation is not worthwhile.

With further development, the framework developed in this article could help city staff address a variety of urban planning and system design problems. For example, several U.S. cities are developing indicator and data collection systems to measure benefits of green infrastructure investments across multiple environmental systems and jurisdictions. The approach in this article should help planners and policymakers conceptualize data design and support decision-making in this heterogeneous physical and institutional context.

ACKNOWLEDGMENTS

We would like to thank our interview participants for their time and generosity in sharing their experiences in smart city-related projects across the Denver metropolitan region, and for helping to connect our research team with additional experts to enrich our investigation. This research was supported by a Dean's Graduate Innovation Assistantship, awarded to the first author by the College of Engineering & Applied Science, University of Colorado Boulder.

REFERENCES

[1] Paul Aoki, Allison Woodruff, Baladitya Yellapragada, and Wesley Willett. 2017. Environmental protection and agency: Motivations, capacity, and goals in participatory sensing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 3138–3150. https://doi.org/10.1145/3025453.3025667

[2] Julio Aráuz. 2018. Smart cities and the dire need for a course correction. In *Proceedings of the 2018 IEEE International Smart Cities Conference (ISC2 2018)*. IEEE, Piscataway, NJ, USA, 1–6. https://doi.org/10.1109/ISC2.2018.8656829

[3] Mariam Asad, Christopher A. Le Dantec, Becky Nielsen, and Kate Diedrick. 2017. Creating a sociotechnical API: Designing city-scale community engagement. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 2295–2306. https://doi.org/10.1145/3025453.3025963

[4] Örjan Bodin. 2017. Collaborative environmental governance: Achieving collective action in social-ecological systems. *Science* 357, 6352 (2017), eaan1114. https://doi.org/10.1126/science.aan1114

[5] danah boyd and Kate Crawford. 2012. Critical questions for big data. *Information, Communication & Society* 15, 5 (2012), 662–679. https://doi.org/10.1080/1369118X.2012.678878

[6] Jeffrey A. Burke, Deborah Estrin, Mark Hansen, Andrew Parker, Nithya Ramanathan, Sasank Reddy, and Mani B. Srivastava. 2006. Participatory sensing. Position paper for the 2006 Workshop on World-Sensor-Web (WSW): Mobile Device Centric Sensor Networks and Applications, co-located with the ACM SenSys 2006 Conference, Boulder, CO. http://citeseer.ist.psu.edu/viewdoc/download?doi=10.1.1.122.3024&rep=rep1&type=pdf

[7] Luís Carvalho. 2014. Smart cities from scratch? A socio-technical perspective. *Cambridge Journal of Regions, Economy and Society* 8, 1 (06 2014), 43–60. https://doi.org/10.1093/cjres/rsu010

[8] Everton Cavalcante, Nélio Cacho, Frederico Lopes, and Thais Batista. 2017. Challenges to the development of smart city systems: A system-of-systems view. In *Proceedings of the 31st Brazilian Symposium on Software Engineering (SBES '17)*. ACM, New York, 244–249. https://doi.org/10.1145/3131151.3131189

[9] City and County of Denver. 2020. Denver Smart City. https://www.denvergov.org/content/denvergov/en/denver-smart-city.html.

[10] City and County of Denver. 2020. Love My Air. https://www.denvergov.org/content/denvergov/en/environmental-health/our-divisions/environmental-quality/air-quality/love-my-air.html.

[11] Sandy Claes, Jorgos Coenen, and Andrew Vande Moere. 2018. Conveying a civic issue through data via spatially distributed public visualization and polling displays. In *Proceedings of the 10th Nordic Conference on Human-Computer Interaction (NordiCHI '18)*. ACM, New York, NY, USA, 597–608. https://doi.org/10.1145/3240167.3240206

[12] Stephen J. Clement, David Wesley McKee, and Jie Xu. 2017. Service-oriented reference architecture for smart cities. In *2017 IEEE symposium on service-oriented system engineering (SOSE)*. IEEE Computer Society, Los Alamitos, CA, 81–85. https://doi.org/10.1109/SOSE.2017.29

[13] Colorado Smart Cities Alliance. 2019. Colorado Smart Cities Alliance. https://coloradosmart.city. Accessed: 2020-05-15.

[14] Eric Corbett and Christopher A. Le Dantec. 2018. Exploring trust in digital civics. In *Proceedings of the 2018 Designing Interactive Systems Conference (DIS '18)*. ACM, New York, NY, USA, 9–20. https://doi.org/10.1145/3196709.3196715

[15] Saskia Coulson, Mel Woods, Michelle Scott, Drew Hemment, and Mara Balestrini. 2018. Stop the noise! Enhancing meaningfulness in participatory sensing with community level indicators. In *Proceedings of the 2018 Designing Interactive Systems Conference (DIS '18)*. ACM, New York, NY, USA, 1183–1192. https://doi.org/10.1145/3196709.3196762

[16] Carl DiSalvo, Tom Jenkins, and Thomas Lodato. 2016. Designing speculative civics. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 4979–4990. https://doi.org/10.1145/2858036.2858505

[17] Rong Du, Paolo Santi, Ming Xiao, Athanasios V. Vasilakos, and Carlo Fischione. 2019. The sensable city: A survey on the deployment and management for smart city monitoring. *IEEE Communications Surveys Tutorials* 21, 2 (Secondquarter 2019), 1533–1560. https://doi.org/10.1109/COMST.2018.2881008

[18] W. Keith Edwards, Mark W. Newman, and Erika Shehan Poole. 2010. The infrastructure problem in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, 423–432. https://doi.org/10.1145/1753326.1753390

[19] Laura Forlano. 2016. Decentering the human in the design of collaborative cities. *Design Issues* 32, 3 (2016), 42–54. https://doi.org/10.1162/DESI{_}a{_}00398 arXiv:https://doi.org/10.1162/DESI_a_00398

[20] Marcus Foth, Irina Anastasiu, Monique Mann, and Peta Mitchell. 2020. From automation to autonomy: Technological sovereignty for better data care in smart cities. In *Automating Cities: Design, Construction, Operation and Future Impact*, Brydon T. Wang and C. M. Wang (Eds.). Springer, Singapore, 319–343.

[21] Guo Freeman, Jeffrey Bardzell, Shaowen Bardzell, Szu-Yu (Cyn) Liu, Xi Lu, and Diandian Cao. 2019. Smart and fermented cities: An approach to placemaking in urban informatics. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 44, 13 pages. https://doi.org/10.1145/3290605.3300274
[22] Daniel Gooch, Matthew Barker, Lorraine Hudson, Ryan Kelly, Gerd Kortuem, Janet Van Der Linden, Marian Petre, Rebecca Brown, Anna Klis-Davies, Hannah Forbes, Jessica Mackinnon, Robbie Macpherson, and Clare Walton. 2018. Amplifying quiet voices: Challenges and opportunities for participatory design at an urban scale. *ACM Transactions on Computer-Human Interaction (ToCHI)* 25, 1, Article 2 (Jan. 2018), 34 pages. https://doi.org/10.1145/3139398
[23] Colin Harrison and Ian Abbott Donnelly. 2011. A theory of smart cities. In *Proceedings of the 55th Annual Meeting of the ISSS*. International Society for the Systems Sciences, Pocklington, UK, 521–535. http://journals.isss.org/index.php/proceedings55th/article/view/1703
[24] Sara Heitlinger, Nick Bryan-Kinns, and Rob Comber. 2019. The right to the sustainable smart city. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 287, 13 pages. https://doi.org/10.1145/3290605.3300517
[25] IEEE. 2020. IEEE International Smart Cities Conference (ISC2 2019). https://ieee-isc2.org/.
[26] Rob Kitchin. 2014. The real-time city? Big data and smart urbanism. *GeoJournal* 79, 1 (2014), 1–14. https://doi.org/10.2139/ssrn.2289141
[27] Rob Kitchin and Niamh Moore-Cherry. 2020. Fragmented governance, the urban data ecosystem and smart city-regions: The case of Metropolitan Boston. *Regional Studies* (2020), 11. https://doi.org/10.1080/00343404.2020.1735627
[28] Stacey Kuznetsov, George Davis, Jian Cheung, and Eric Paulos. 2011. Ceci n'est pas une pipe bombe: Authoring urban landscapes with air quality sensors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 2375–2384. https://doi.org/10.1145/1978942.1979290
[29] Stacey Kuznetsov and Eric Paulos. 2010. Participatory sensing in public spaces: Activating urban surfaces with sensor probes. In *Proceedings of the 8th ACM Conference on Designing Interactive Systems (DIS '10)*. ACM, New York, NY, USA, 21–30. https://doi.org/10.1145/1858171.1858175
[30] Christopher A. Le Dantec and W. Keith Edwards. 2010. Across boundaries of influence and accountability: The multiple scales of public sector information systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 113–122. https://doi.org/10.1145/1753326.1753345
[31] John Lee and Neville Moray. 1992. Trust, control strategies and allocation of function in human-machine systems. *Ergonomics* 35, 10 (1992), 1243–1270. https://doi.org/10.1080/00140139208967392
[32] Monique Mann, Peta Mitchell, Marcus Foth, and Irina Anastasiu. 2020. #BlockSidewalk to Barcelona: Technological sovereignty and the social license to operate smart cities. *Journal of the Association for Information Science and Technology (JASIST)* 71, 9 (2020), 1103–1115. https://doi.org/10.1002/asi.24387
[33] Donald McMillan, Arvid Engström, Airi Lampinen, and Barry Brown. 2016. Data and the city. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 2933–2944. https://doi.org/10.1145/2858036.2858434
[34] Alfredo Mela. 2014. Urban public space between fragmentation, control and conflict. *City, Territory and Architecture* 1, 1 (2014), 1–7. https://doi.org/10.1186/s40410-014-0015-0
[35] Microsoft. 2019. Denver's pioneering Smart City program improves public safety, mobility, and health. https://customers.microsoft.com/en-us/story/city-and-county-denver-government-azure.
[36] Mile High Flood District. 2020. Mile High Flood District. https://mhfd.org/.
[37] Matthew B.. Miles and A Michael Huberman. 1994. *Qualitative data analysis: An expanded sourcebook*. SAGE Publications, Thousand Oaks, CA.
[38] Milind Naphade, Guruduth Banavar, Colin Harrison, Jurij Paraszczak, and Robert Morris. 2011. Smarter cities and their innovation challenges. *Computer* 44, 6 (June 2011), 32–39. https://doi.org/10.1109/MC.2011.187
[39] National Science Foundation. 2019. Smart and Connected Communities. https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505364.
[40] Paolo Neirotti, Alberto De Marco, Anna Corinna Cagliano, Giulio Mangano, and Francesco Scorrano. 2014. Current trends in Smart City initiatives: Some stylised facts. *Cities* 38 (2014), 25–36. https://doi.org/10.1016/j.cities.2013.12.010
[41] Safiya Umoja Noble. 2018. *Algorithms of oppression: How search engines reinforce racism*. NYU Press, New York.
[42] Patrick Olivier and Peter Wright. 2015. Digital civics: Taking a local turn. *interactions* 22, 4 (July–August 2015), 61–63. https://doi.org/10.1145/2776885
[43] Cathy O'Neil. 2016. *Weapons of math destruction: How big data increases inequality and threatens democracy*. Broadway Books, New York.
[44] Open Mobility Foundation. 2020. Mobility Data Specification. https://github.com/openmobilityfoundation/mobility-data-specification.
[45] Open Mobility Foundation. 2020. Open Mobility Foundation. https://www.openmobilityfoundation.org/.

[46] Krassimira Paskaleva, James Evans, Christopher Martin, Trond Linjordet, Dujuan Yang, and Andrew Karvonen. 2017. Data governance in the sustainable smart city. *Informatics.* 4, 4 (November 2017), 41. http://dro.dur.ac.uk/24121/
[47] Peña Station NEXT. 2020. Peña Station NEXT. https://penastationnext.com/.
[48] Riccardo Petrolo, Valeria Loscri, and Nathalie Mitton. 2017. Towards a smart city based on cloud of things, a survey on the smart city vision and paradigms. *Transactions on Emerging Telecommunications Technologies* 28, 1 (2017), e2931. https://doi.org/10.1002/ett.2931
[49] Joseph Pierce, Deborah G. Martin, and James T. Murphy. 2011. Relational place-making: The networked politics of place. *Transactions of the Institute of British Geographers* 36, 1 (2011), 54–70. https://doi.org/10.1111/j.1475-5661.2010.00411.x
[50] Aare Puussaar, Ian G. Johnson, Kyle Montague, Philip James, and Peter Wright. 2018. Making open data work for civic advocacy. *Proceedings of the ACM on Human-Computer Interaction (PACM HCI)* 2, CSCW, Article 143 (Nov. 2018), 20 pages. https://doi.org/10.1145/3274412
[51] Ricky Robinson, Markus Rittenbruch, Marcus Foth, Daniel Filonik, and Stephen Viller. 2012. Street computing: Towards an integrated open data application programming interface (API) for cities. *Journal of Urban Technology* 19, 2 (2012), 1–23. https://doi.org/10.1080/10630732.2012.698064
[52] Jathan Sadowski. 2020. *Too smart: How digital capitalism is extracting data, controlling our lives, and taking over the world.* MIT Press, Cambridge, MA.
[53] Johnny Saldaña. 2015. *The coding manual for qualitative researchers.* SAGE Publications, Thousand Oaks, CA.
[54] Steve Sawyer and Mohammad Hossein Jarrahi. 2014. The sociotechnical perspective. In *CRC Handbook of Computing (Vol. 5(1))*, A. Tucker and H. Topi (Eds.). Chapman and Hall, New York, 5–27.
[55] Taylor Shelton, Matthew Zook, and Alan Wiig. 2014. The 'actually existing smart city'. *Cambridge Journal of Regions, Economy and Society* 8, 1 (10 2014), 13–25. https://doi.org/10.1093/cjres/rsu026
[56] Bhagya Nathali Silva, Murad Khan, and Kijun Han. 2018. Towards sustainable smart cities: A review of trends, architectures, components, and open challenges in smart cities. *Sustainable Cities and Society* 38 (2018), 697–713. https://doi.org/10.1016/j.scs.2018.01.053
[57] Susan Leigh Star. 1999. The ethnography of infrastructure. *American Behavioral Scientist* 43, 3 (Nov. 1999), 377–391. https://doi.org/10.1177/00027649921955326
[58] Susan Leigh Star and James R. Griesemer. 1989. Institutional ecology, 'translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science* 19, 3 (Aug. 1989), 387–420. https://doi.org/10.1177/030631289019003001
[59] Anselm Strauss and Juliet Corbin. 1990. *Basics of qualitative research.* SAGE publications, Thousand Oaks, CA.
[60] Alex S. Taylor, Siân Lindley, Tim Regan, David Sweeney, Vasillis Vlachokyriakos, Lillie Grainger, and Jessica Lingel. 2015. Data-in-place: Thinking through the relations between data and community. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15).* ACM, New York, NY, USA, 2863–2872. https://doi.org/10.1145/2702123.2702558
[61] U.S. Department of Transportation. 2016. Smart City Challenge. https://www.transportation.gov/smartcity.
[62] Nitya Verma and Lynn Dombrowski. 2018. Confronting social criticisms: Challenges when adopting data-driven policing strategies. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18).* ACM, New York, NY, USA, Article 469, 13 pages. https://doi.org/10.1145/3173574.3174043
[63] Amy Voida, Ellie Harmon, and Ban Al-ani. 2011. Homebrew databases: Complexities of everyday information management in nonprofit organizations. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI '11).* ACM Press, New York, 915–924. https://doi.org/10.1145/1978942.1979078
[64] Alan Wiig. 2016. The empty rhetoric of the smart city: From digital inclusion to economic promotion in Philadelphia. *Urban Geography* 37, 4 (2016), 535–553. https://doi.org/10.1080/02723638.2015.1065686
[65] Langdon Winner. 1980. Do artifacts have politics? *Daedalus* 109, 1 (1980), 121–136. http://www.jstor.org/stable/20024652
[66] Tan Yigitcanlar, Marcus Foth, and Md. Kamruzzaman. 2019. Towards post-anthropocentric cities: Reconceptualizing smart cities to evade urban ecocide. *Journal of Urban Technology* 26, 2 (2019), 147–152. https://doi.org/10.1080/10630732.2018.1524249
[67] Shoshana Zuboff. 2019. *The age of surveillance capitalism.* PublicAffairs, New York.

Received June 2020; revised October 2020; accepted December 2020