

# A feeling for the data: How government and nonprofit stakeholders negotiate value conflicts in data science approaches to ending homelessness

Stephen C. Slota<sup>1</sup>  | Kenneth R. Fleischmann<sup>1</sup>  | Min Kyung Lee<sup>1</sup> |  
 Sherri R. Greenberg<sup>2</sup> | Ishan Nigam<sup>1</sup> | Tara Zimmerman<sup>1</sup> |  
 Sarah Rodriguez<sup>3</sup> | James Snow<sup>4</sup>

<sup>1</sup>School of Information, The University of Texas at Austin, Austin, Texas, USA

<sup>2</sup>LBJ School of Public Affairs, University of Texas at Austin, Austin, Texas, USA

<sup>3</sup>Department of Design and Delivery, City of Austin, Austin, TX, USA

<sup>4</sup>Department of Public Works, City of Austin, Austin, TX, USA

## Correspondence

Stephen C. Slota, School of Information, The University of Texas at Austin, Austin, Texas, USA.

Email: [steveslota@gmail.com](mailto:steveslota@gmail.com)

## Funding information

Good Systems

## Abstract

Governmental and organizational policy increasingly claims to be data-driven, data-informed, or knowledge-driven. We explore the data practices of local governments and nonprofits seeking to end homelessness in the City of Austin. Drawing on 31 interviews with stakeholders, alongside the reflections and experiences of our interdisciplinary, cross-sector collaborative team, we consider the role of data in guiding and informing interventions and policy regarding homelessness. Ending homelessness is a particularly challenging scenario for intervention, with increasing politicization, changing circumstances, and needing rapid intervention to reduce harm. In exploring some implications of data science “in the wild” as it is deployed, understood, and supported within the Travis County Continuum of Care (CoC), we analyze how data-intensive work connects and engages across disciplinary boundaries. Furthermore, we consider how data science and the iField can collaborate in addressing complex, social problems as advisors and partners with invested organizations.

## 1 | INTRODUCTION

Applying data science requires not only an intimate understanding of data science methods, but also, of the data itself. In her profile of geneticist Barbara McClintock, famous for her discovery of genetic transposition, (Ravindran, 2012), historian of science Evelyn Fox Keller writes,

“What is it in an individual scientist's relation to nature that facilitates the kind of seeing that eventually leads to productive discourse? What enabled McClintock to see further and deeper into the mysteries of genetics than her colleagues? Her answer is simple. Over and over again, she tells us one

must have the time to look, the patience to ‘hear what the material has to say to you,’ the openness to ‘let it come to you.’ Above all, one must have ‘a feeling for the organism.’” (Keller, 1983, pp. 197–198).

Scholars in the information field (iField) have identified examples of why data scientists must similarly develop “a feeling for the data.” Having “a feeling for the data” means having a nuanced understanding and appreciation for what the data is, where it came from, and its strengths and limitations. Having the data and working with the data is not enough, one must understand the history and life of the data, in a deep sense. Operating on data provided by others, and viewing it and operating upon the data uncritically, is unlikely to result in “a feeling for the

data.” Gaining this perspective can be best facilitated by witnessing the data collection process and interacting with those who collect the data to understand the factors that influence how they interact with and manipulate the data. For example, Marathe and Toyama (2021) describe how epilepsy diagnosis is socially constructed through decisions of which events to consider as seizures, which may have implications for diagnosis and treatment. Doctors may have a particular diagnosis or treatment regime in mind, and their sense of their intended outcome may influence how they choose to document or to fail to document seizures. They argue, “we propose that the successful adoption of [machine learning] and [clinical decision support systems] in healthcare requires a contextual understanding of clinicians’ work” (p. 13). Similarly, Anderson et al. (2019), argue that social media posts about disasters should not only be understood in the aggregate, but also, that each post should be considered within its own context. The goal of this paper is to explore how government and nonprofit stakeholders develop “a feeling for the data” and negotiate value conflicts as they work to end homelessness.

The concept of data has changed over time, and in line with knowledge productive practices. Rosenberg (2013) documents how the notion of data shifted from that which is known to the results of experimentation and focused inquiry in the 18th century, to further describing any digital objects that might be manipulated by a computer. Bowker (2005) notes a further conceptual shift of data closely tied to scientific memory practices and the technological infrastructure supporting that work. Hence, a modern understanding of data might be closer to any record of observations retrieved electronically from a computer. Metaphorically, data might be understood as something to be consumed, refined, or controlled as a side effect of our behavior over computational infrastructures, including social and consumptive behaviors—a “resource” that shortens the link between evidence and action (Puschmann & Burgess, 2014). However, we are reminded that “raw data is an oxymoron” (Gitelman, 2013), and the processes of storing, describing, and otherwise ordering information are tied to consequential categories and classifications (Bowker & Star, 2000), with metadata (data about data) and relevant standards serving as a bridge to interoperability between disparate datasets (Millerand & Bowker, 2009). Therefore, we can understand “data” as structured set of digitally-stored observations of that world that are structured and interpreted to negotiate between the unstructured “world” and knowledge infrastructures (Edwards et al., 2013), informed by practice and bearing social, discursive, and economic consequence.

Calls for increasingly “datafied” (Lycett, 2013) approaches to governance emphasize the potential power of data to inform decision-making, to maximize resource use, and “to transform the design and delivery of public policies and services in ways that improve social goods and impact societal wellbeing” (van Ooijen et al., 2019, p. 53). Similarly, “smart cities” built upon data collection and analysis regularly envision a utopian environment of resident engagement and knowledge-informed policy (Ismagilova et al., 2019; Kurzweil & Grossman, 2010). These approaches often fail to fully consider the importance of context in analyzing data.

Much as information technology has reduced the impact of geographical distance on the diffusion of culture (Castells, 2010; Graham, 1998) and sociality (Laniado et al., 2018), governance has become increasingly diffuse. Governance can no longer be seen as solely the actions of the state in administration and policy implementation (Lynn et al., 2000), but as an expansive theory considering “lateral relations, interinstitutional relations, the decline of sovereignty, the diminishing importance of jurisdictional borders, and a general institutional fragmentation” (Frederickson & Smith, 2003, p. 226). Yang (2012) argues that “involvement of non-governmental individuals or organizations in public administration and public affairs challenges traditional state-centered management and forms a new social organization and management system” (p. 586).

The challenges of governance through, and with, data are more expansive than just doing good data analysis. Aside from organizational and institutional factors like data silos, inconsistent description, and access restrictions, data-driven governance grapples with justice, sociality, and responsiveness to changes in politics and circumstances. We are considering the initial phase of a collaborative project addressing the ongoing management of homelessness within the Travis County Continuum of Care (CoC). We relate qualitative results from our study and account for how they will inform the development of a data-oriented intervention. We highlight the need for a range of expertise to begin addressing this “wicked problem” (Buchanan, 1992; Head, 2008), and we conclude with how the strengths of the iField can supplement and engage with data science in developing knowledge-driven, policy interventions.

## 2 | BACKGROUND

Trust is a key factor in effective, data-driven responses to crises, especially when the response requires decentralized engagement and public participation (Petridou, 2020). The ideal form of this trust might be summarized as “the

public sector provides reliable information and evidence for the citizens to make informed decisions while the citizens in turn are trusted to conduct themselves responsibly” (Petridou, 2020, p. 154), but that ideal, naturally, never quite exists. Data-driven governance can create a level of “distance” between government action and the lived experience of those being governed, limiting opportunities for resistance or shifting them into novel forms specific to the technologies and institutions involved (Smith & O’Malley, 2017). Dencik et al. (2019) argue that data collection and analysis of a city’s individual residents shifts the boundaries of individual and collective risks, and often results in an inability for residents to challenge the results of analysis.

Aspects of governance can be automated, where data drawn from sensor networks supplants traditional processes of government, such as the automation of speed enforcement (Smith & O’Malley, 2017). This is even more troubling when differential treatment results from data-intensive tools and analyses: “When it comes to the so-called data revolution the power is firmly in the hands of those who are able to interpret or tell stories with the data,” (Beer, 2018). To use data, however, you must understand the data’s context and its flaws, to develop “a feeling for the data.” Data, depending on the services provided, may exist in silos and need to be centrally provided as a service to data scientists to produce meaningful insights (von Wachter et al., 2020). Further, analysis may be limited by lack of counterfactual evidence rendering it impossible to understand the causal effects of the automated tools developed by data scientists (Kube et al., 2019).

Chelmiss et al. (2021) call for a close relationship between holders of domain knowledge and data scientists when approaching homelessness service provision, and they call for site-specific data scientific research, closely informed by the needs of the served community and service providers. While the importance of domain knowledge to data scientific intervention has long been known (Ribes et al., 2019), the specifics of what knowledge should be necessary for that work are not yet well-understood. Even assessing whether a given dataset could be useful for providing insight—prior to performing any analysis—represents significant knowledge work (Slota et al., 2020). This represents a pressing opportunity to further explore how implicit knowledge, existing practice, or organizational arrangements contribute to needed domain knowledge in pursuing data scientific interventions. According to Yang (2012), scholarly engagement in governance projects typically improves their outcomes, especially when there is little political intervention throughout the project’s life.

However, shifting policy landscapes have the potential to radically change goals, measures, or the overall

environment—a dynamic that can create significant tension as the rhythms of data work become in tension with those of the environment of that work (Jackson et al., 2011). The strength of data-driven predictions can be improved where close knowledge of specific service users complements the knowledge that can be obtained from analyzing population-level datasets (Vaithianathan & Kithulgoda, 2020). Further, collecting sub-population specific data while involving domain experts and services users can generate relevant data-driven insights that may shape public policies to improve the quality of life of sheltered service users (Hong et al., 2018).

Personal meetings may represent the main communications channel for those experiencing homelessness, and refinement of the technological and systemic enablers of those meetings can be productive in engaging with potential services and increasing agency among this community (Johansson & Gulliksen, 2019). Social media, despite the challenges in maintaining access to communications technology (Woelfer & Hendry, 2011), can be a key means by which many people experiencing homelessness engage with services and each other (Koepler, 2014), but one that cannot be assumed among any segment of the population due to lack of consistent access to devices and infrastructure. When people on the homelessness continuum have access to technology, such as mobile phones, they become vital tools by which this community ensures their safety (Hendry et al., 2011) and finds employment (Hendry et al., 2017). In considering how to design services, however, we must remember that these tools are not universally present, or even consistently available. Being able to pay for data and plans, having consistent access to infrastructure like electrical outlets for charging, and a higher risk of theft all are barriers to engagement. When designing services for people experiencing homelessness, technological solutions that are not undergirded by in-person, or other processes that do not require a specific device or internet access, exclude many in the population. Thus, the personal motivations and values of people managing homelessness become a central factor in ensuring adoption and engagement with new tools and shared insight (Fitzpatrick & Stephens, 2014).

In a study of social media behaviors of people experiencing homelessness, Koepler et al. (2013) present a series of “values portraits” through which values conflicts and synergies among stakeholders in systems design might be more effectively understood. Even in systems yet to be built, algorithmic imaginaries (Ames, 2018) of potential systems reflect the values and needs of their designers, and can be effectively informed through stakeholder engagement, through structured interventions to elicit perspectives on values and social contexts of systems use (Yoo et al., 2013). Similarly, exploring design

issues through human values can help analyze and assess existing information systems for people experiencing homelessness (Burrows et al., 2019). Value-sensitive design has informed the exploration of Twitter posts by people experiencing homeless, identifying their statistical likelihood to value helpfulness and wealth more than people not experiencing homelessness (Koepler & Fleischmann, 2012), as well as how those values structure engagement with technology (Woelfer et al., 2011). Le Dantec et al. (2009) argue for a more empirical, rather than discursive, approach to value-sensitive design that emphasizes discovering values through lived experience. This approach emphasizes continued reflections on values, rather than entrenching a specific ethical philosophy of values.

There are relatively few studies that consider housing access for under-resourced or marginalized populations through information, communication, and technology (Shamsuddin & Srinivasan, 2021), representing a significant opportunity for research, especially when combined with studying how data and ICT are used to inform decision making. Similarly, an improved understanding of how personal and organizational values impact the understanding and use of data represents a key opportunity for better coordinating knowledge-driven insight with practice in ending homelessness.

### 3 | METHODS

The following research questions guided the themes we present here, and the analysis of those themes in the context of this writing:

1. How do government and nonprofit stakeholders resolve value conflicts related to data about people experiencing homelessness?
2. How can data professionals balance values that impact access to data (e.g., privacy)?
3. How can data science and other disciplines mutually support each other in creating data—or knowledge-driven policy?

#### 3.1 | Research site

This study relates the initial qualitative findings of a long-term, mixed-method engagement with the City of Austin. The City of Austin collaborates with a CoC, a regional organization that coordinates access to federal assistance funding. Led by a dedicated organization known as ECHO (Ending Community Homelessness Coalition) the CoC coordinates and shares resources with

the heterogeneous providers of services, shelter, or other resources in the area (ECHO, 2021). These service providers vary in mission, approach, and funding sources. Within the CoC, information resources are coordinated through an HMIS (Homelessness Management Information System) database managed by ECHO. This approach, along with the CoC organizational model, is recommended by the US Department of Housing and Urban Development (HUD) and is a common approach to managing urban homelessness across the United States.

The City of Austin has invested substantially in more effectively using its data—a key example is the Affordable Housing Search Tool, which resulted from the 2019 Code for America Summit, a collaboration between the City of Austin, the Housing Authority of the City of Austin, Code for America, Austin CityUP, the Austin Tenants Council, one of the co-authors of this paper, and a representative from the tech industry (City of Austin, 2020). This work, which began in 2017, produced a tool to facilitate finding affordable housing, and it coordinated data from the City of Austin, the Housing Authority of the City of Austin, the Housing Authority of Travis County, and the Texas Department of Housing and Community Affairs. This project builds on the City of Austin's prior work in making better use of its data. The CoC uses various data-collection tools to prioritize service to people experiencing homelessness, most notably coordinated assessment of vulnerability through a tool known as the VI-SPDAT (Slota et al., 2021). Our collaboration seeks to develop knowledge-driven, AI-based, interventions and insights. First, we explored the data, collaborations, and organizational dynamics of stakeholders within the CoC. In presenting these findings, we draw upon the case study model (Yin, 2012) to demonstrate the results of an open-ended, initial, research phase.

#### 3.2 | Participants

We interviewed 31 individuals engaged with service provision or information for people on the homelessness continuum. Interviewees were selected following recommendations from key informants from the City of Austin and a co-author of this study, who were collaborators on this project, then further developed through snowball sampling, to leverage the unique social knowledge of participants (Noy, 2008). Nineteen of the interviewed participants held leadership roles within their organizations (Manager, Director, or Coordinator titles), and the remainder held more direct service provision roles, such as working on homelessness “street teams” to identify areas of service need or providing contract and grant

support. Of these participants, 4 were from ECHO, 14 from nonprofits, 12 from collaborating local or state governments, and 1 participant who was from an academic entity.

### 3.3 | Data collection and analysis

To draw deeply from the lived experience of individuals directly engaged with the knowledge infrastructure of the CoC, we structured these interviews according to critical incident technique (Flanagan, 1954). In these interviews, we addressed our participants data needs and practices, and the outcome and output measures that they used or understood by each participant. Interview protocols were initially drafted by the first two authors, then discussed and refined with the full team, including key collaborators from the City of Austin. We asked each interviewee to relate three recent projects or events related to their work with people experiencing homelessness: one that was not yet resolved, one that resolved in a particularly successful or satisfying manner, and one that resolved in a less-satisfying or more frustrating manner. For each critical incident, we asked questions about the background of the project, existing and desired data for the project, existing and desired collaborations for the project, and reflections on the outcomes of the project. Through this, we sought to ground responses in recent, specific, experiences and interactions, and to elicit both tensions and synergies in collaborative data work practices among participants.

We then conducted a series of practice interviews with volunteers in roles similar to those we would be formally interviewing. We discussed these practice interviews and the feedback from our initial volunteers with the full team, and further refined the instrument towards our goal of understanding and representing data and collaborative practices. Interviews were conducted by four different members of our research team, operating in pairs, and following each interview we debriefed on potential themes, interesting topics, and potential refinements of our data co-creation process. As interviews were conducted, we discussed the progress of interviews and our impressions of early findings during weekly team meetings, and identified further potential participants, in line with snowball sampling techniques. This was a key step in understanding and characterizing the institutional environment for our team and an aid in directing potential interventions. Interviews took approximately 1 h and were recorded over teleconferencing software.

We analyzed the transcribed interviews inductively using thematic analysis (Braun & Clarke, 2006), and identifying key themes in the data. We selected thematic

analysis due to its versatility in describing rich qualitative datasets. In early reviews of the data, the two lead authors identified descriptions of various projects, their outcomes, and the resources that made them possible, as well as barriers to success. These broad categories became our initial codes, which guided our further iterative rounds of coding. The lead authors sorted and coded the data, applying additional codes and subcodes as needed. Next, we reviewed and assessed the codes and subcodes to generate overall themes, which we reviewed to ensure that they accurately represented the dataset. The research team, as a whole, agreed to organize and present the data using these themes, including descriptive quotes for each.

Our open coding process resulted in a series of themes, some of which are presented in this paper. One theme focused on the collaborative practices of our interviewed participants, leading us to conduct an additional, deductive, round of coding focusing on how values guided systemic approaches to ending homelessness within the Continuum of Care (Zimmerman et al., 2021). Another set of themes revolved around the impact of time and temporality of the viability, utility, and efficacy of data collected to support service provision, and is the topic of a forthcoming conference paper (Slota et al., 2022). These themes, while relevant to our understanding of the site and important in developing our interpretation of further data, did not adequately serve to inform the next planned phase of our team's overall project, and its intended data scientific intervention. To do this, we returned to a consequential instrument that served as both a point of contact for data collection and a means of prioritizing services based on prior research regarding risks to health and safety, which we had briefly discussed in terms of infrastructural justice (Slota et al., 2021).

In later rounds of coding and theme development, we revisited this theme in terms of how it works to describe the relationship between service provision, internal and external data, and the influence of prior research in operationalizing a concept of risk in prioritizing service provision. This refined theme, along additional themes related to justice, data sharing, and systemic/social visibility, are presented in this paper. In exploring these themes, and with an eye towards our larger research team's upcoming data scientific interventions, we approached this paper considering what it takes to build upon existing successes in the approaches we had previously observed through our interviews. As we worked to better understand how we might contribute to the successes of the CoC, our research team became sensitized to both the role of data science in establishing given conceptions of risk, enabling systemic and social visibility of the community of people experiencing homelessness, and introducing novel

collaborative and values tensions in organizing ongoing work. It is in this light that we began to analyze collected qualitative data towards answering the questions posed above.

## 4 | FINDINGS

### 4.1 | The VI-SPDAT and data-oriented construction of risk

The Continuum of Care makes use of a coordinated assessment tool to prioritize service provision and limited resources across the community. In our interviews, we found that this assessment process, alongside annual counts, was both a vital tool in “seeing” the community and understanding their needs. The specific tool used to perform these coordinated assessments, the VI-SPDAT (Vulnerability Index—Service Prioritization Decision Assistance Tool) makes use of risk factors identified in medical research conducted by Boston’s Healthcare for the Homeless program (Hwang et al., 1997), structured by a decision support tool developed by OrgCode Consulting, Inc., to assess socioeconomic, medical, and psychological risk factors of homelessness. This tool plays a vital role in the how homelessness is understood in the CoC and bears significant individual consequence in its effective construction of the concept of risk within the community. However, this tool was recognized in the Continuum of Care as being significantly flawed, and its revision and improvement was a point of emphasis among many of the CoC stakeholders, as described by the below participant, who played a leadership role in the assessment process:

Past our pilot, [we have] implemented additional equity points on top of the existing scoring system ... And basically, the plan moving forward along with our HUD support is to just kind of make sure that we’re monitoring our outcomes and the kind of assessment pieces on a very regular scale, to make sure that we are prioritizing folks early on for housing resources, and that we can kind of right this inequity that we see in the disproportionate overrepresentation of Black African-Americans in our homeless population.

Arising from both medical research and a more human-centered approach to deploying data to guide decision-making, the VI-SPDAT represents both the distant action of data science in supporting policy and operations “on

the ground” and a key means of observing and understanding the served community, even though its flaws were well known, as described below.

And, you know, with that coordinated entry system, they do an intake, and they do, you know, an acuity assessment, and a score, whatever. And that’s how they’re prioritized. But, you know, I appreciate the system, because I think it does create some transparency, and also accountability ... but I also think there’s still this population that just isn’t able to do that assessment sometimes, or because they’re so ill, and they don’t realize they have issues or they’re very paranoid. So, they don’t want to tell anybody their business. They may score lower acuity when they’re actually some of the most high-acuity folks.

What we saw in these accounts was not the uncritical acceptance of the VI-SPDAT as a tool, nor strict reliance upon data collected through its use. Rather, our participants recognized the need to localize a general tool to the needs of their community, and in line with the values they held in terms of its outcomes. Experience with the use of the VI-SPDAT and ongoing engagement with people who were represented by the data collected through it, led to adjustments both in how data were understood, and in how collection happened. Lived experience, and the tacit knowledge that came with it, was consistently related as vital to making things happen:

A lot of the people that we serve are so ill that they were not scoring as a priority through the continuum of care process, they were scoring very low as a matter of fact, probably because they’re not the best advocates for themselves, given their mental health and substance use issues. So, they were never going to get housing through the continuum of care, while we were still seeing the severe need for them to be housed, because they were literally dying on the street ... The reason they’re not getting prioritized is because those coordinated assessments are based on self-report.

Through this, we can point to a feeling, not just for the data, but for how that data worked systemically to represent, prioritize, and facilitate access to services for the community, especially in contrast to how users advocate for or represent themselves. The VI-SPDAT, though a

vital tool in both representing the community of people experiencing homelessness and serving as a means of prioritizing access to services, was responsible for only a piece of the data collected across the CoC. Point in Time counts, records of services being offered, on-site observations, and even records of watershed cleanup and use of food pantries are all collected, with heterogeneous levels of coordination and interoperability across the CoC, towards the goal of understanding and providing services. However, the VI-SPDAT, unlike many of the other data sources, represented a point of coordination across all service providers, and as such was the focus of much of the conversation about data quality and utility.

## 4.2 | Data silos

The most common organizational structure we encountered relegated data management, analysis, and sharing to relatively few individuals—often only a single individual in the organization due to licensing costs. These individuals served as key points of articulation in coordinating the use and sharing of data among the myriad service providers, including nonprofit organizations, and local government entities. As data management was concentrated in the hands of relatively few individuals, individuals in leadership roles in each organization that we interviewed characterized their data as remote, and insufficient. Yet, the data was consistently characterized as vital, as a need and a means of progressing their work: “And yeah, that data sharing, I cannot overemphasize that. Because we use it in real time to inform the work and improve the work.” Participants consistently related a desire for more data, better coordinated data, or more comprehensive inter-organizational knowledge sharing, believing that the current offerings were insufficient. Data silos were consistently problematized. Considering how these silos might be broken down prompted reflection by some participants as to how the knowledge gained through their work might benefit the broader CoC. This information had few formal channels through which it could be shared but numerous informal, relationship-driven, channels. Even where data was considered “solved” in the context of that organization’s work, they recognized that they were able to provide insights for other organizations engaged in the CoC, as stated below:

So in terms of data that I wish we could collect? I think we solved a lot of that at this point. Other than I’d really like to see this as an opportunity to get people back into the HMIS system or find ways to assess in the

field when we’re delivering their food, because oftentimes, we’re the only people that they’re in contact with a week and figuring out what is their health situation.

In the above quote, this participant acknowledges that their data—sufficient to their own purposes and effective for internal needs—had uses outside of that context but lacked pathways for sharing and external use. This participant imagined that the HMIS system could potentially serve as such a pathway but acknowledged that this information currently was not being shared. Participants, overall, recognized that data sharing was only one aspect of the knowledge sharing necessary for coordination. In this case, data sharing, including the breaking of data silos, was merely one tactic leveraged towards the overall strategy of inter-organizational knowledge sharing and coordination. Participants recognized that specialized knowledge was vital, but that the broader goals of data sharing (inter-organizational knowledge sharing and coordination) could be achieved through a variety of levels of engagement with data.

## 4.3 | Building a social impact model: How to sustain data sharing

In this section, we present a brief narrative related by a participant in a leadership role within HMIS, describing an effort to engage in productive data sharing, “social impact modeling.” Social impact modeling, in the case of this participant, referred to an effort to better understand how money spent on homelessness service impacted the state of homelessness within the CoC, with an overall goal of demonstrating how money is saved or lost through their permanent supportive housing policy. This collaboration was one of the largest in recent history of the City of Austin, and engaged City offices with Travis County offices, local criminal justice data (i.e., court records, arrest records, etc.), healthcare data managed through Central Health, and the United States Department of Housing and Urban Development. Of note is the initial complexity of managing any sort of data sharing agreement—from a legal standpoint especially—in grappling with local organizational policy and data management practice. The social impact model that was the initial goal of the collaboration became, for this participant, primarily an issue of negotiating data sharing: “the end game was ... this really big data sharing agreement between many entities. So we had that guidance. And that was really helpful.”

Thus, we see the data sharing agreement becoming the apparent “end game” in the collaboration, even

though it was intended initially, and primarily, to produce a model that would allow more effective resource prioritization, as described below:

Where we could really clearly say, you know, if I house Bri, based on ... a brief history in our healthcare system, and our criminal justice system, and even our ... shelter system, that ... Bri cost our community X amount of money ... and that by providing the intervention of housing and supportive services, we can now clearly demonstrate from data from the back end how much money essentially as being saved as a system.

However, even with collected data, and a broad data-sharing agreement, something was still missing from the project. The social impact model was not achieved, and there was some discomfort in replacing what had previously been done by humans through experience and tacit knowledge with a data scientific approach. Where humans had previously been responsible for prioritizing efforts and evaluating outcomes, there was some discomfort in relegating that work to data analysis:

That felt really strange to pursue, because it just took away ... a lot of the human element of this. But additionally, we just couldn't figure that out. Um, yeah. And so in some ways, it was a failure.

The outcomes of this project did not match its initial goals:

"... we received data, and were able to do some data matching in our system, which is typically, if we have a choice we prefer to do is to receive a match on our own ... we just weren't able to identify a way to do that, **at a large scale on an ongoing basis.**" (emphasis added)

"Data matching" involves ensuring that one set of records refers to the same individual as another set of records. While data matching was a common goal of this specific data sharing initiative, unfortunately, it saw limited success. In this case, the infeasibility of the desired social impact model was not solely a result of the characteristics of the data, its analysis, or exclusively of the organizational dynamics and temporal concerns of participating groups, but rather, it emerged at the intersection of these issues. "I'd like to believe there was a way we could have managed the relationships better ... that we could

have fostered more buy in." The high bar for data sharing of legally protected data, alongside gaps in the data itself and organizational dynamics defining how data are useful, impeded the sustainable collaboration that would have been needed to achieve the desired social impact model. Both social and financial factors were mentioned as sustainability gaps, though the social factors (development of ongoing partnerships or continual, internally motivated, work in managing data sharing, etc.) were front-lined. However, there were still valuable results of the project:

... The project [is] still moving forward. It just looks very different. So we did ... create space for some really significant investments, especially from the city of Austin, and some foundations. So we did create that relationship and buy-in in some spaces, and that part is moving forward. It'll end up being, as of now, one of the largest permanent supportive housing projects we've put on the ground as a community in quite some time ...

What was created, instead of sustainably de-siloed data or an effective social impact model, were more durable relationships, oriented towards the goal of seeking funding, that led to closer collaboration and the development of effective ongoing projects somewhat extraneous to the initial goal. In terms of providing research-driven knowledge and interventions, there would be a need to recognize outcomes that were not originally planned to be measured, and that might be resistant to understanding through singular measures.

#### 4.4 | Systemic and social visibility

Beyond outcome measurement and refinement to internal processes, participants also related emerging tensions from the notion of visibility, broadly conceived. In the case of these interviews, we identify systemic invisibility, where individuals receiving services do not have consistently coordinated records across systems or are otherwise lost in data collection and analytic practices, as a key tension in effective service provision. Existing data collection regimes, individual assessment, and organizational policy often had gaps where significantly vulnerable individuals would fall, as with the below individual who was incapable of advocating for themselves, but had no guardian to act on their behalf:

And it wasn't honestly anything that a homeless service provider could like, have offered

... It was like he needed a capacity assessment and guardianship. But there wasn't a tool for that. And so it's kind of like how do you like what do you do with ... people that are sort of still invisible to me, if that makes sense?

There was no process by which the client referenced above could receive the guardianship and capacity assessment needed for the provider to advocate and work on their behalf. This systemic invisibility might be hinted at within the data collected, but only comes to the forefront through personal experience, exists primarily as tacit knowledge, and should be addressed through policy and process revisions rather than through analytic tools. As participants consistently valued broad coverage and universal access to services above and beyond efficiency, it becomes a challenge to account for the people who fall through the gaps in existing processes. Participants related significant extra effort in attempting to address people who become invisible, as the service provider below relates:

Clients, when they go missing, we have a process for that ... how we look for them, and what to do if we don't find them ... [which] was so needed ... we had clients going missing or falling out of contact very regularly, which is to be expected ... [as] they don't have a safe and stable place to live yet. They're having to move around every day for their basic needs ... [and] keeping in touch with the case manager is, it's not an immediate and basic need for them.

This challenge—keeping track of users who become broken or lost within collected data—was of concern across different service providers, and in city offices. Visibility, in this circumstance, might be achieved through better or more consistent data practice, but was most commonly addressed through organizational policy, and the creation of processes to find and account for people who become lost after entering the CoC system. Thus, we see a key tension in undertaking data work to inform policy. Available data was ineffective in tracing and accounting for many individuals. This resulted in a loss of visibility for a significant portion of the population. Participants related visibility to be important as community response was seen to be a major driver of policy and funding support for service provision:

The same thing kind of happened for us with COVID. When we moved outside ... we were

... serving a similar amount of people, pre-pandemic, as we are kind of right now. But people are just like, holy cow, you guys are serving so many people and really, it's because we're making them stand six feet apart on a sidewalk versus having them all inside. So it's just more of a visibility question than it is [a question of] are we helping more people? I don't know, maybe a ... few more but like, a lot of it has to do with visibility. And so visibility often times drives response.

In the above quote, the invisibility arose, not from tracking, quantification, or data traces, but from the simple fact of being indoors rather than outdoors. This form of social invisibility was perceived to have a direct impact on policy formation and public response, one that often could shift policy or funding in unpredictable ways, further complicating predictive analysis of policy approaches. Many participants felt that the scope of homelessness in their area was often underestimated by people not directly engaged in the work, and that such invisibility (and the related visibility when individuals “come outside”) drives action, both positive and negative, as stated below:

... And I think there's a misconception in the community that the removal of that [camping] ban somehow quadrupled our population. And that's not at all what happened. We just saw them as more visible right, because they felt safer and allowed to camp in places that were closer to service provision, or were closer to natural resources that they needed, like clean water and the ability to access transportation. And unfortunately, the narrative became that we just allow homelessness to explode across the city.

According to the above participant, little changed after rescinding the camping ban other than the location of people experiencing homelessness, but that change prompted a discursive shift in the communities impacted by that change. Where above we saw visibility driving positive, often altruistic, engagement, it could also drive negative discursive shifts resulting in less friendly policies, or less support for ongoing work delivering primarily long-term results. And for these last two quotes, it is important to remember that mostly they describe people who are already accounted for in the data yet remain socially invisible.

In this theme, visibility plays two complementary roles: the first oriented towards the inter-organizational knowledge necessary to adequately offer services over time to the same individuals, the second towards representing the state and capacity of the system to a broader public. Both aspects of visibility were at play, and both were concerned with the viability and utility of the data collected in support of the work of service provision. While being able to effectively represent local conditions was a vital aspect of this work in terms of gaining access to funding, supporting policy initiatives, and driving charitable contributions, being “lost” in the system resulted in more immediate harms to service users due to delays or inconsistencies in delivering support. Adequate representation in data across the system, then, works to empower both the organization itself and the people being served by it. However, as we discuss in the next section, the issue of agency over personal data among people served by the system becomes further complicated as disclosure becomes linked with health, safety, and access to housing.

#### 4.5 | Agency and self-advocacy

Visibility regarding how data is used, and how individuals might be able to express some agency over use, also become key factors in making data-informed decisions, especially when prioritizing limited resources. Participants reported that their clients were often limited in understanding how the data collected about them might be used, which can lead to resistance in providing that data, and limits access to services made available only to people who score above a threshold. Resistance to data collection, as shared by the following interviewee who worked for the City of Austin, was related as being engendered by a lack of understanding of how personal information was deployed in prioritizing access to services.

I had a client who ... she had not done her assessment so there was nothing in HMIS [and I had to] scrape and peel for months just to get where ... I could actually help her. So I feel like ... if she understood why the ECHO assessment was important, why getting her information at HMIS was important. I think that would help ... there's really like, no real explanation in the frontline, to say, this is why you're doing this. And this is how it's going to help you.

This represents a key loss of data agency for people seeking access to services. People seeking services were, as part of the standard assessment process, asked to self-

report on areas of their life that they would otherwise not want to share. The standard initial assessment, which seeks to measure vulnerability and provides a key initial point of contact in providing services, includes questions about drug use and abuse, medical history, and family dynamics. While the need to navigate the tension between effective data collection and privacy is immediately apparent in this case, this tension is characteristic of the loss of data agency and information that occurs as data is consolidated, de-siloed, or shared across service providers. This is significant when people using the system are expected, to some extent, to effectively advocate for themselves to obtain access to support and services. In the quote below, a nonprofit service provider relates a frustration held with encouraging a person to advocate for themselves according to the terms of the data collection instrument used to inform service prioritization.

We provide the coordinated assessment and we have to ask it and utilize it the way that they expect us to ... [like with this individual who] we have been serving for many years that comes in and out of the system. This individual did have intellectual development, developmental disabilities and had been in and out of prison and jail for his behavior. When asked, ‘do you have access to food?’ He said yes. And his access to food was eating out of the dumpster. And we knew that because we would see it.

Embarrassment, desire to protect one's privacy, and the fear of admitting to illegal behavior were all related as barriers to effective self-advocacy for people seeking services. Navigating these instruments effectively would require intimate knowledge of their operation and how their results are deployed and would potentially result in significant personal discomfort when disclosing the personal information needed to do so, with cascading effects throughout the system as these assessments are a key piece of data collected consistently across participating services in the CoC.

We see within this theme a series of broken chains of reference and inscription in tracing, accounting for, and representing the state of homelessness and people experiencing homelessness. People who cannot self-advocate for reasons of mental health or disability are in a disadvantaged position in initial assessment of vulnerability. People are often lost within the system due to their lack of permanent address and/or access to the internet. Knowledge about people experiencing homelessness as an aggregate was not readily available to the community, and when that knowledge comes about as people are

physically confronted with the realities of homelessness, it can be interpreted by that community as ineffective processes. Again, a commingling of concerns about data and its collection with concerns related to contextualizing and representing that information to community members and policy makers. When working in an area so dependent on voluntary contributions and appeals to grant funders, this representation becomes a pivotal factor in participants' organizations' ability to achieve their goals and desired outcomes.

## 5 | DISCUSSION

Our participants related that much of their problem-solving and daily work was enabled and supported through personal relationships, tacit knowledge, and other deeply social means, in line with Johansson and Gulliksen's (2019) findings that emphasize the importance of personal meetings in providing services to people experiencing homelessness. Similarly, we saw some of the effects of Dencik et al.'s (2019) assertion that data-intensive regimes provide little leverage to contest the outcomes of algorithmic governance. This limits the ability of service users to exercise personal agency over their data during assessment processes. This also follows Petridou's (2020) assertion of the need for trust between the community and government in crisis response as distrust could be a key factor in unwillingness to share personal information that might lead to a higher priority for service provision. While we saw evidence of benevolent negotiation of the system, in line with Marathe and Toyama (2021), where assessor's experience and familiarity with the assessment process enabled immediate intervention to ensure more accurate depictions of risk within the system. This level of familiarity was the result of long-term engagement with the community. Furthermore, there are significant complications in the creation, use, and management of data specific to homelessness.

Our findings echo a finding from Neff et al. (2017) that communication is a critical component of data science, and we should consider "the entire process of data gathering and production as one that has communication at its core and communication practices as key data science practices." Complementing emerging work that investigates technical processes of data science (Passi & Jackson, 2018; Zhang et al., 2020), such as data cleaning and model training, our findings show that the field of data science can leverage the theories and methods from iFields to understand and design sustainable and ethical data science practices and data-informed policy. The management, curation, and sharing of information has long been a strength of the iField. Many data scientific

concerns, such as those raised by van Ooijen et al. (2019), deal closely with issues of domain knowledge, context, and curation—an arena where the iField is particularly well-equipped from its history and current research. Prior work offers examples of research methods (e.g., understanding human values of AI engineers [Shilton et al., 2014]) and design frameworks and methods (e.g., participatory frameworks to involve stakeholders in AI design [Lee et al., 2019]) that we can use to create common grounds among those who collect, analyze and use data and provides ways to preserve context. Similarly, issues of the just use of data and effective self-advocacy, which become significantly magnified in the context of data about people experiencing homelessness (Eubanks, 2018), can only be partially addressed through algorithmic refinement and better data—they require adjustments to practice and policy. Even good data can still produce unjust outcomes (Eubanks, 2018), and where data are flawed or unrepresentative, this can only introduce additional concerns.

In addressing an issue of significant social concern, knowledge-informed policy requires coordination and significant effort beyond the more commonly anticipated issues of data quality and access. Complicating factors include broader, sustainable inter-organizational coordination, management of how the environment and individuals within it were perceived and represented (both through data collected and knowledge held by individuals), and even how individuals within the system understood how their data might be used. These factors affect the implementation of policy changes and the perceived, and actual, quality of the data used to inform action. Similarly, decision-making and daily case management occurred not only through information systems, but also, through non-obvious social relationships built upon individual and organizational histories. In the account of the CoC present in these interviews, we see a system that is heterogeneous in work and data practice. It is one that is often ad hoc as services respond to changing circumstances while valuing rapid intervention, and subject to shifting levels of inter-organizational coordination over time.

This study was solely focused on a single CoC within the United States. While it is likely that in other policy regimes the level of agency over personal data ensured by policy (as with the GDPR in Europe) may create additional options for people experiencing homelessness, there is still a level of knowledge about policy and relevant systems necessary to do so. Much as our service providers related accounts of their users lacking the capacity to effectively advocate for themselves, there is reason to believe that available options for expressing agency over how users are represented and accounted for in data might be similarly remote in other policy regimes.

However, the GDPR may make individuals more comfortable with or likely to share personal and medical information (Karampela et al., 2019), which could enable more accurate assessment.

This study helps to illustrate how a human-centered data science approach could transform work in the public sector. To effectively incorporate data science into social service practice, it is critical both for data scientists to have a feeling for the data and for social service providers to have a feeling for the algorithm (Stevens, 2017). As Ribes et al. (2019) argue, data scientists need to have domain knowledge relevant to the application area from which the data comes and in which the results of the analysis will be applied. Limitations to timeliness and interoperability of data have significant implications for how data can be used as well as for how to interpret the analyses produced through a data science approach. Similarly, understanding population-level dynamics might be useful in seeking and apportioning resources, but seemed to bear little impact on the daily activities of social service providers, which were commonly related by our participants as being reactive to changing circumstances that are not well-represented in data resources. Where data science can have a significant impact on this site, then, is in the refinement of data collection practices, like through analysis and refinement of the VI-SPDAT process according to its population-level outcomes, and through tools that enable reasoning and tracking across the diversity of data collected across organizations. In considering the results of this study, we emphasize the role of domain knowledge and close accounts of practice in guiding data scientific intervention, with attentiveness to the specific needs of the site in identifying where those interventions might have the greatest impact.

There are things that an algorithm can “see” that a human cannot, and things a human can “see” that remain invisible to the algorithm and data. It is when these visions are coordinated that impactful intervention becomes possible. Data scientists need to make their algorithms as transparent and explainable as possible, so that users of the products of data science analysis such as social service providers can understand the built-in assumptions and limitations of the data and how it is algorithmically interpreted. The iField is the ideal meeting place for human-centered data science practice and user-centered design of interfaces that leverage data science techniques. As such, the findings of this study help to illustrate the potential harms that can be caused by the use of data science without consideration of its context, and the ethical imperative to use data-science in a human-centered way that grounds interpretation of the data and algorithm in terms of how they will be applied in the real world. As such, the iField can help to ensure

that data science can be leveraged towards building a more equitable and just society.

In seeking to understand how practitioners used data in informing, enabling, and assessing their services, we found an intensely social, values-driven, mode of work across a complex ecology of organizational histories and individual relationships driving the collection, sharing, and use of data. Long-term assessments of the overall population being served can become occluded by high levels of focus on the particulars of daily work. Without accounting for those particulars, however, data can become decontextualized or siloed and actionable analyses become increasingly distant to the practical concerns. Effective knowledge-driven policy, then, can be seen through these interviews as negotiating immediacy and personal experience with trends and dynamics that become visible only over time. Hence, interventions can easily become an inter-organizational, intensely political negotiation of the values and priorities of individuals engaged in service provision. Computational reasoning and computationally derived knowledge become useful in directing work, prioritizing resources, and assessing outcomes when paired with a deep understanding of the social, organizational, and ethical dynamics of stakeholders in the system—a process that might be thought of as bringing data “to the ground” of specific contexts and circumstances.

## 6 | CONCLUSION

While data science promises insight and a novel means of obtaining value and guidance for policy, deployment of those insights necessarily engages with the daily work of practitioners with and without data scientific expertise. As governments seek to leverage large volumes of previously disconnected data, concerns of context, local practice, and short-term reactivity become increasingly vital in contextualizing data-driven insight.

Our research findings inform the practice of data science and the design of its surrounding system in the context of services for people experiencing homelessness. For example, it calls for an effort to make people who are currently invisible to be rendered more visible in data. The optimization or prediction goals of data science should aim for not only understanding the risk of being homelessness based on the VI-SPDAT data but whose risk is not being captured by the VI-SPDAT system. Using consolidated, de-siloed data sources and stakeholder consultation, one would need to determine an appropriate alternative indicator of risk for such populations. It will be also important to create a system that supports for different stakeholders, whether they are social workers, the

leadership and management teams, data scientists, and the public; their differing roles are all crucial in effective service provision, whether it is about making better service recommendations, budget allocation decisions, or keeping the data up-to-date. Our work provides insights into their varying information needs, and highlights that the importance of visualizations and narratives of analytics results, and annotations and documentation of data collection in supporting them.

As our work progresses, we will begin qualitative data collection among people experiencing homelessness, and we will more fully account for how service users understand data produced about them, and the impacts of data on their interactions with service providers. Additionally, in collaboration with local government offices and service providers, we are working towards knowledge-driven interventions in service provision and prioritization through data science methods and informed by our qualitative research. The findings presented here provide key insight for future studies that engage with how reliance on computational analysis might isolate or marginalize people who are represented in that data, and how they might advocate for themselves or express agency over that data.

“Real world” problems and “real world” data are rarely neat, organized, complete, and consistent. The iField can provide perspective on the context of data and its social construction. Hence, useful data scientific interventions become environments where the tools and perspectives of the iField contextualize, represent, and understand the domains with which they work (Ribes et al., 2019; Slota et al., 2020). Training in the iField can thus produce human-centered data scientists who achieve “a feeling for the data.”

## ACKNOWLEDGMENTS

This work was supported by Good Systems, A UT Grand Challenge #Y2P11: “Smart Cities Should Be Good Cities: AI, Equity, and Homelessness.” We wish to thank our collaborators Michelle Surka, Keyanna S. Evans, Khalil Bholat, Divya Rathanlal, and Jonathan Tomko for their organizational, methodological, and conceptual contributions to this project. Thanks also to all of our study participants, including those who chose to be named and credited for their intellectual contributions to this study: Patricia Barrera, Bill Brice, Veronica Buitron Camacho, Samantha Campbell, Sarah Garvey, Shontell Gauthier, Alex Graham, David Gomes, Mark Hilbelink, Andy Hoffmeister, Vella Karmen, Robert Kingham, Audrey Kuang, Nora Linares-Moeller, Allison Mabbs, Susan McDowell, Walter Moreau, Jo Kathryn Quinn, Katie Rose, Lisa Ruiz, Julia Spann, Whitney Thurman, Pete Valdez, Soleece Watson, Bree Williams, and the seven participants who chose to remain anonymous.

## ORCID

Stephen C. Slota  <https://orcid.org/0000-0001-6526-8030>

Kenneth R. Fleischmann  <https://orcid.org/0000-0002-0323-3526>

## REFERENCES

- Ames, M. G. (2018). Deconstructing the algorithmic sublime. *Big Data & Society*, 5(1), 1–4. <https://doi.org/10.1177/2053951718779194>
- Anderson, J., Casas Saez, G., Anderson, K., Palen, L., & Morss, R. (2019). Incorporating context and location into social media analysis: A scalable, cloud-based approach for more powerful data science. In *Proceedings of the 52nd Hawaii international conference on system sciences*, IEEE Computer Society Press, Hawaii, 2274–2283. <http://hdl.handle.net/10125/59666>
- Beer, D. (2018). Envisioning the power of data analytics. *Information, Communication & Society*, 21(3), 465–479. <https://doi.org/10.1080/1369118X.2017.1289232>
- Bowker, G. C. (2005). *Memory practices in the sciences*. MIT Press.
- Bowker, G. C., & Star, S. L. (2000). *Sorting things out: Classification and its consequences*. MIT Press.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp0630a>
- Buchanan, R. (1992). Wicked problems in design thinking. *Design Issues*, 8(2), 5–21. <https://doi.org/10.2307/1511637>
- Burrows, R., Mendoza, A., Sterling, L., Miller, T., & Pedell, S. (2019). Evaluating ask izzzy: A mobile web app for people experiencing homelessness. In *Proceedings of 17th European conference on computer-supported cooperative work*. European Society for Socially Embedded Technologies (EUSSET). <https://dl.eusset.eu/handle/20.500.12015/3250>
- Castells, M. (2010). Globalisation, networking, urbanisation: Reflections on the spatial dynamics of the information age. *Urban Studies*, 47(13), 2737–2745. <https://doi.org/10.1177/0042098010377365>
- Chelms, C., Qi, W., & Lee, W. (2021, April). Challenges and opportunities in using data science for homelessness service provision. In *Companion proceedings of the web conference 2021*, Association for Computing Machinery, New York, 128–135. <https://doi.org/10.1145/3442442.3453454>
- City of Austin. (2020). *Austin announces new online affordable housing search tool (AHOST)*. Press Release. <http://www.austintexas.gov/news/austin-announces-new-online-affordable-housing-search-tool-ahost>
- Dencik, L., Redden, J., Hintz, A., & Warne, H. (2019). The ‘golden view’: Data-driven governance in the scoring society. *Internet Policy Review*, 8(2), 1–24. <https://doi.org/10.14763/2019.2.1413>
- ECHO. (2021). *Continuum of care*. <https://www.austinecho.org/leadingsystem-change/continuum-of-care/>
- Edwards, P. N., Jackson, S. J., Chalmers, M. K., Bowker, G. C., Borgman, C. L., Ribes, D., Burton, M., & Calvert, S. (2013). *Knowledge infrastructures: Intellectual frameworks and research challenges*. Deep Blue. <http://hdl.handle.net/2027.42/97552>
- Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.
- Fitzpatrick, S., & Stephens, M. (2014). Welfare regimes, social values and homelessness: Comparing responses to marginalised

- groups in six European countries. *Housing Studies*, 29(2), 215–234. <https://doi.org/10.1080/02673037.2014.848265>
- Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin*, 51(4), 327–358. <https://doi.org/10.1037/h0061470>
- Frederickson, H. G., & Smith, K. B. (2003). *The public administration theory primer*. Westview Press.
- Gitelman, L. (Ed.). (2013). *Raw data is an oxymoron*. MIT Press.
- Graham, S. (1998). The end of geography or the explosion of place? Conceptualizing space, place and information technology. *Progress in Human Geography*, 22(2), 165–185. <https://doi.org/10.1191%2F030913298671334137>
- Head, B. W. (2008). Wicked problems in public policy. *Public Policy*, 3(2), 101–118.
- Hendry, D. G., Abokhodair, N., Kinsley, R. P., & Woelfer, J. P. (2017). Homeless young people, jobs, and a future vision: Community members' perceptions of the job co-op. In *Proceedings of the 8th international conference on communities and technologies*, Association for Computing Machinery, New York, 22–31.
- Hendry, D. G., Woelfer, J. P., Harper, R., Bauer, T., Fitzer, B., & Champagne, M. (2011). How to integrate digital media into a drop-in for homeless young people for deepening relationships between youth and adults. *Children and Youth Services Review*, 33(5), 774–782.
- Hong, B., Malik, A., Lundquist, J., Bellach, I., & Kontokosta, C. E. (2018). Applications of machine learning methods to predict readmission and length-of-stay for homeless families: The case of win shelters in New York City. *Journal of Technology in Human Services*, 36(1), 89–104. <https://doi.org/10.1080/15228835.2017.1418703>
- Hwang, S. W., Orav, E. J., O'Connell, J. J., Lebow, J. M., & Brennan, T. A. (1997). Causes of death in homeless adults in Boston. *Annals of Internal Medicine*, 126(8), 625–628. <https://doi.org/10.7326/0003-4819-126-8-199704150-00007>
- Ismailova, E., Hughes, L., Dwivedi, Y. K., & Raman, K. R. (2019). Smart cities: Advances in research—An information systems perspective. *International Journal of Information Management*, 47, 88–100. <https://doi.org/10.1016/j.ijinfomgt.2019.01.004>
- Jackson, S. J., Ribes, D., Buyuktur, A., & Bowker, G. C. (2011). Collaborative rhythm: Temporal dissonance and alignment in collaborative scientific work. In *Proceedings of the ACM 2011 conference on computer supported cooperative work*, Association for Computing Machinery, New York (pp. 245–254). <https://doi.org/10.1145/1958824.1958861>
- Johansson, S., & Gulliksen, J. (2019). *Can mainstream smart technology support homeless people leaving homelessness?*. <http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-262816>
- Karampela, M., Ouhbi, S., & Isomursu, M. (2019). Exploring users' willingness to share their health and personal data under the prism of the new GDPR: Implications in healthcare. In *2019 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, IEEE. 6509–6512. <https://doi.org/10.1109/EMBC.2019.8856550>
- Keller, E. F. (1983). *A feeling for the organism: The life and work of Barbara McClintock*. W.H. Freeman.
- Koepfler, J. A. (2014). *Values and self-presentation in online communication by stakeholders related to homelessness* (Unpublished doctoral dissertation). University of Maryland, Baltimore, MD. <https://www.proquest.com/docview/1561147445>
- Koepfler, J. A., & Fleischmann, K. R. (2012). Studying the values of hard-to-reach populations: Content analysis of tweets by the 21st century homeless. In *Proceedings of the 2012 iConference*, Association for Computing Machinery, New York, 48–55. <https://doi.org/10.1145/2132176.2132183>
- Koepfler, J. A., Shilton, K., & Fleischmann, K. R. (2013). A stake in the issue of homelessness: Identifying values of interest for design in online communities. In *Proceedings of the 6th international conference on communities and technologies*, Association for Computing Machinery, New York, 36–45. <https://doi.org/10.1145/2482991.2482994>
- Kube, A., Das, S., & Fowler, P. J. (2019). Allocating interventions based on predicted outcomes: A case study on homelessness services. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(1), 622–629. <https://doi.org/10.1609/aaai.v33i01.3301622>
- Kurzweil, R., & Grossman, T. (2010). Bridges to life. In G. M. Fahy, D. M. D. West, L. S. Coles, & S. B. Harris (Eds.), *The future of aging* (pp. 3–22). Springer.
- Laniado, D., Volkovich, Y., Scellato, S., Mascolo, C., & Kaltenbrunner, A. (2018). The impact of geographic distance on online social interactions. *Information Systems Frontiers*, 20(6), 1203–1218. <https://doi.org/10.1007/s10796-017-9784-9>
- Le Dantec, C. A., Poole, E. S., & Wyche, S. P. (2009). Values as lived experience: Evolving value sensitive design in support of value discovery. In *Proceedings of the SIGCHI conference on human factors in computing systems*, Association for Computing Machinery, New York, 1141–1150. <https://doi.org/10.1145/1518701.1518875>
- Lee, M. K., Kusbit, D., Kahng, A., Kim, J. T., Yuan, X., Chan, A., See, D., Noothigattu, R., Lee, S., Psomas, A., & Procaccia, A. D. (2019). WeBuildAI: Participatory framework for algorithmic governance. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), 1–35. <https://doi.org/10.1145/3359283>
- Lycett, M. (2013). 'Datafication': Making sense of (big) data in a complex world. *European Journal of Information Systems*, 22(4), 381–386. <https://doi.org/10.1057/ejis.2013.10>
- Lynn, L. E., Jr., Hcinrich, C. J., & Hill, C. J. (2000). Studying governance and public management: Why? How? In C. J. Heinrich & L. E. Lynn, Jr. (Eds.), *Governance and performance: New perspectives* (pp. 1–33). Georgetown University Press.
- Marathe, M., & Toyama, K. (2021). The situated, relational, and evolving nature of epilepsy diagnosis. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW3), 1–18. <https://doi.org/10.1145/3432916>
- Millerand, F., & Bowker, G. (2009). Metadata standards: Trajectories and enactment in the life of an ontology. In M. Lampland & S. L. Star (Eds.), *Standards and their stories: How quantifying, classifying, and formalizing practices shape everyday life* (pp. 149–165). Cornell University Press.
- Neff, G., Tanweer, A., Fiore-Gartland, B., & Osburn, L. (2017). Critique and contribute: A practice-based framework for improving critical data studies and data science. *Big Data*, 5(2), 85–97. <https://doi.org/10.1089/big.2016.0050>
- Noy, C. (2008). Sampling knowledge: The hermeneutics of snowball sampling in qualitative research. *International Journal of Social Research Methodology*, 11(4), 327–344. <https://doi.org/10.1080/13645570701401305>

- Passi, S., & Jackson, S. J. (2018). Trust in data science: Collaboration, translation, and accountability in corporate data science projects. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW), 1–28. <https://doi.org/10.1145/3274405>
- Petridou, E. (2020). Politics and administration in times of crisis: Explaining the Swedish response to the COVID-19 crisis. *European Policy Analysis*, 6(2), 147–158. <https://doi.org/10.1002/epa2.1095>
- Puschmann, C., & Burgess, J. (2014). Big data, big questions: Metaphors of big data. *International Journal of Communication*, 8, 1690–1709.
- Ravindran, S. (2012). Barbara McClintock and the discovery of jumping genes. *Proceedings of the National Academy of Sciences*, 109(50), 20198–20199. <https://doi.org/10.1073/pnas.1219372109>
- Ribes, D., Hoffman, A. S., Slota, S. C., & Bowker, G. C. (2019). The logic of domains. *Social Studies of Science*, 49(3), 281–309. <https://doi.org/10.1177%2F0306312719849709>
- Rosenberg, D. (2013). Data before the fact. In L. Gitelman (Ed.), “Raw data” is an oxymoron (pp. 15–40). MIT Press.
- Shamsuddin, S., & Srinivasan, S. (2021). Just smart or just and smart cities? Assessing the literature on housing and information and communication technology. *Housing Policy Debate*, 31(1), 127–150. <https://doi.org/10.1080/10511482.2020.1719181>
- Shilton, K., Koepfler, J. A., & Fleischmann, K. R. (2014). How to see values in social computing: Methods for studying values dimensions. In *Proceedings of the 17th ACM conference on computer supported cooperative work & social computing*, Association for Computing Machinery, New York, 426–435. <https://doi.org/10.1145/2531602.2531625>
- Slota, S. C., Fleischmann, K., & Greenberg, S. (2022). A brief typology of time: Temporal structuring and dissonance in service provision for people experiencing homelessness. In M. Smits (Ed.), *Information for a better world: Shaping the global future. iConference 2022. Lecture notes in computer science* (Vol. 13192, pp. 211–224). Springer. [https://doi.org/10.1007/978-3-030-96957-8\\_19](https://doi.org/10.1007/978-3-030-96957-8_19)
- Slota, S. C., Fleischmann, K., Greenberg, S., Surka, M., Evans, K. S., Snow, J., Rodriguez, S., & Zimmerman, T. (2021). Just infrastructure? Field research on a standardized assessment tool for a continuum of care for people experiencing homelessness. In *Proceedings of the 84th annual meeting of the Association for Information Science and Technology (ASIS&T)*, 58(1), 327–336. <http://doi.org/10.1002/pra2.460>
- Slota, S. C., Hoffman, A. S., Ribes, D., & Bowker, G. C. (2020). Prospecting (in) the data sciences. *Big Data & Society*, 7(1), 205395172090684. <https://doi.org/10.1177%2F2053951720906849>
- Smith, G. J., & O'Malley, P. (2017). Driving politics: Data-driven governance and resistance. *The British Journal of Criminology*, 57(2), 275–298.
- Stevens, H. (2017). A feeling for the algorithm: Working knowledge and big data in biology. *Osiris*, 32(1), 151–174. <https://doi.org/10.1086/693516>
- Vaithianathan, R., & Kithulgoda, C. I. (2020). *Using predictive risk modeling to prioritize services for people experiencing homelessness in Allegheny County: Methodology update*. Auckland University of Technology, Centre for Social Data Analytics.
- van Ooijen, C., Ubaldi, B., & Welby, B. (2019). A data-driven public sector: Enabling the strategic use of data for productive, inclusive and trustworthy governance. In *OECD Working Papers on Public Governance*. No. 33. OECD Publishing. <https://doi.org/10.1787/09ab162c-en>
- von Wachter, T., Rountree, J., Buenaventura, M., Blackwell, B., & Obermark, D. (2020). *Evaluation of Los Angeles county measure H-funded homelessness prevention strategies*. UC Berkeley, California Policy Lab. <https://escholarship.org/uc/item/0s67t6gt>
- Woelfer, J. P., & Hendry, D. G. (2011). Homeless young people and technology: Ordinary interactions, extraordinary circumstances. *Interactions*, 18(6), 70–73. <https://doi.org/10.1145/2029976.2029994>
- Woelfer, J. P., Iverson, A., Hendry, D. G., Friedman, B., & Gill, B. T. (2011). Improving the safety of homeless young people with mobile phones: Values, form and function. In *Proceedings of the SIGCHI conference on human factors in computing systems*, Association for Computing Machinery, New York, 1707–1716.
- Yang, L. (2012). Building a knowledge-driven society: Scholar participation and governance in large public works projects. *Management and Organization Review*, 8(3), 585–607. <https://doi.org/10.1111/j.1740-8784.2011.00279.x>
- Yin, R. (2012). *Applications of case study research* (3rd ed.). Sage Publications.
- Yoo, D., Huldtgren, A., Woelfer, J. P., Hendry, D. G., & Friedman, B. (2013). A value sensitive action-reflection model: Evolving a co-design space with stakeholder and designer prompts. In *Proceedings of the SIGCHI conference on human factors in computing systems*, Association for Computing Machinery, New York, 419–428. <https://doi.org/10.1145/2470654.2470715>
- Zhang, A. X., Muller, M., & Wang, D. (2020). How do data science workers collaborate? Roles, workflows, and tools. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW1), 1–23. <https://doi.org/10.1145/3392826>
- Zimmerman, T., Slota, S., Fleischmann, K. R., Greenberg, S., Snow, J., & Rodriguez, S. (2021). Socio-technical collaboration to end homelessness: A case study of perspectives on the value of interorganizational cooperation for data sharing. In *TMS proceedings*. <https://tmb.apaopen.org/pub/4fx9rcsq>

**How to cite this article:** Slota, S. C., Fleischmann, K. R., Lee, M. K., Greenberg, S. R., Nigam, I., Zimmerman, T., Rodriguez, S., & Snow, J. (2022). A feeling for the data: How government and nonprofit stakeholders negotiate value conflicts in data science approaches to ending homelessness. *Journal of the Association for Information Science and Technology*, 1–15. <https://doi.org/10.1002/asi.24715>