

PLAY #4

QUESTION DATA

Data is at the heart of smart cities, as data is what makes technology “smart.” They provide the computational values about the world that technology calculates and activates. The real-time analysis of smart city data can lead to better service provision and information accessibility; however, it can also lead to profiling and social isolation. **Public sector and public-serving organizations need to ask why, how, and what of data before it is collected and interpreted.** Robust and responsible data policies and protocols must be included in the development of smart technology infrastructure, and organizations should make these policies and protocols known to the publics they serve.

should be to question the use of new datasets and to clearly communicate the values that justify their collection and analysis. “We have a burden — as practitioners, as researchers, as policymakers — to make as strong an attempt as we can at explaining the effects and explaining the principles of data use. We need to strive for interpretability,” says Julia Stoyanovich. This is brought into focus by Steve Walter: “If data is owned by government, it is necessarily everyone’s data.” As a result, government needs to think through the consequences of such ownership. Pedestrian traffic data being collected because it serves the purpose of timing walk signals would seem to be desirable. But when that same dataset is mined by an advertising company for demographic information, that may be an unwanted and unintended consequence.

DISCUSSION



Kathy Nyland cautions to not “romanticize data, but to recognize its power and its consequences.” Her statement captures the spirit behind this play: data is often seen as a resource to be controlled, but the public sector’s responsibility is to steward this new data landscape with clear justifications. Nearly everything can be data-fied, but that does not mean it should be collected and stored. Government’s role

Government is not the only actor in public data use; civil society organizations need to play a significant role in stewarding the data landscape. Layman Lee states, “Nothing’s free. Somebody’s collecting our data. If someone’s trying to make money off that data, instead of them going directly to city government on a very wide scale, what happens if the neighborhood owns them?” If a community organization owns the data, they can use them for the purposes they want, without making them available to every organization and company that wants them.

There are significant complications in how data is shared across agencies or organiza-

tions. As Sari Ladin-Sienne states, the City of Los Angeles' job is not simply to make data available; "its job is to educate people on how to contextualize and use data." Cicely Garrett discussed the complexity of sharing sufficiently contextualized data across departments in the City of Atlanta, saying, "There isn't a shared summary of data. I'm not sure that people know a good way to share data."

The responsibility of data collection, management, and analysis within government is often left to understaffed "innovation and technology" offices or designated "smart city" staff. In reality, citizens' digital lives reach across government agencies. To provide more integrated and targeted services for publics, governments must also share and manage data across agencies. Data sharing involves data classification, standardization, and cross-departmental communication, while *informed* data sharing is even more difficult. This involves discussion about which data to share or make public, time spent on providing context to datasets, and attention to the gaps in data collection and analytics. As Adam Forman says, "When sharing open data between departments and with the public, **governments shouldn't adopt a data culture of yes; they should adopt a data culture of why.**"

Gabriel Mugar asks, "Why, if we optimize our personal lives with Fitbits, can't we consider how to optimize civic data, and decide as a community how to use it?" Smart city public engagements are not town hall meetings about capital projects; they are collective processes to determine how to manage data. Symposium participants proposed a flipped data ownership model, such as creating citizen councils for data sharing. As Carl DiSalvo suggests, "Instead of police giving citizens a report, what if citizens gave the city a report about their own data?" A data ombudsman or civic technology jury are options for making these sorts of decisions.

There is too often an assumption that data speaks for itself, and that data cannot be distorted or misinterpreted. Martijn de Waal reflected on a project ran by Het Nieuwe Insti-

tuut in the City of Eindhoven (The Netherlands) that did this work well: "The project started with a lot of existing data, so a lot was known about which problems were at stake in which neighborhoods. But then the organizers went into the communities, asked people to tell their stories, then mapped those stories to the known problems. Turns out that the problems people found relevant were not the ones that local government had data about. For example, people identified loneliness as a problem, but that wasn't represented in any existing data set." Predictive policing is another example: when the majority of arrests in a particular neighborhood involve Black men, an algorithm justifies that Black men be more thoroughly surveilled, thus leading to more arrests.⁹ One response to data bias is context for data sets, as adding context to data can help to surface bias through providing a critical lens on the use of algorithms and analyses, as well as filling in gaps in data.

ACTION IDEAS



- Adopt a data culture of thinking critically about how, when, and why to collect and use data
- Find data about problems, not problems in the data you have
- Empower a civic technology jury to veto smart city technology proposals that don't provide public value

⁹ Ferguson, A. G. (2016). Policing predictive policing. *Washington University Law Review*, 94.