Toward community-inclusive data ecosystems: Challenges and opportunities of open data for community-based organizations

Ayoung Yoon | Andrea Copeland

Abstract

The benefits of open data for helping to address societal problems and strengthen communities are well recognized, and unfortunately previous studies found that smaller communities are often excluded from the current data ecosystem because of existing technological, technical, cognitive, and practical barriers. This study aims to investigate the process of communities’ data use for community development and decision-making—focusing on the opportunities and challenges of data for communities. From the interviews with 25 staff from community-based organizations (CBOs) in nine small, medium, and large cities in the United States, the findings of this study describe data’s role in supporting communities’ development while reporting several major challenges that hinder CBOs data use: difficulty accessing data, limitations of open data (un-local nature, excluding essential data from being open), limited data capacity (especially in data literacy skills), and difficulties using and accessing existing data infrastructures. Our findings suggest opportunities for addressing these challenges, particularly by creating educational programming, building partnerships within data ecosystems, and bringing community voices forward in current data ecosystems, which are critical to realizing data’s potential for all citizens.

1 | INTRODUCTION

The benefits of utilizing open data to address societal problems and strengthen communities are well recognized. Mostly driven by the open government data initiative, open data are generally understood as data that are accessible, machine-readable, free (or available at negligible cost), and have minimal limitations on their use (Bhargava et al., 2015). Many researchers argue that ensuring access to open data has the potential to enhance government accountability and transparency—improving service delivery and public administration while fostering citizens’ participation in civil society (United Nations Development Programme, n.d.; Jofre, Szigeti, & Diamond, 2016; Veeckman, McCrory, & Walravens, 2017). Corruption was one of the main motivations behind the open government data movement, given its power to undermine human rights and economic development. Acts supporting transparency increased participatory governance and access to data need to create social and economic gains (Attard, Orlandi, Scerri, & Auer, 2015). Further, the United Nations Secretary-General’s...
Independent Expert Advisory (IEAG) Group (2014) distinguished the role that data can play in the realization of sustainable development and argued the significance for all stakeholders, such as government, regional institutions, public/private data producers, academics, and citizens, to work together to create a more equitable data ecosystem.

While most previous studies on open data are either conceptual (Bertot, Jaeger, & Grimes, 2010; McDermott, 2010) or focus on open data tool/technology development (Kalampokis, Tambouris, & Tarabanis, 2011), some described the empirical use of open data (Hausenblas, 2009; Napoli & Karaganis, 2010). Still, researchers conduct the majority of studies in the contexts of large cities or at urban levels (Janssen, Charalabidis, & Zuiderwijk, 2012; Martin, Foulonneau, Turki, & Ihadjadene, 2013; Thakuriah, Dirks, & Keita, 2017), as well as developing countries (Aronson, 2004; Bentley & Chib, 2016; Mutuku & Mahihu, 2014; Ntawiha & Anderson, 2016). This means that the current literature fails to demonstrate data’s potential at the local level in smaller communities. Bertot, Butler, and Travis (2014) point out that, in the United States alone, more than 10,000 cities, towns, and villages lack the capacity to engage in the data initiatives that larger cities, national governments, and international nongovernmental organizations (NGOs) pursue, but most studies overlook the fact that many communities operate on a much smaller, local scale. In view of the challenges posed by the local context, it is interesting to note that open data for local development initiatives have been largely funded by international and governmental financial institutions. For example, the World Bank, International Development and Research Centre, and USAID, invested in open data in developing nations because it supports their ongoing commitment to economic development. Another example of the awareness of value of local open data is the World Wide Web Foundation’s support of Open Data Labs (n.d.), which promotes the use of open data for local social and economic benefit. While the importance of local open data is understood, the reach of the global initiatives has not penetrated most communities.

As decisions become more data-driven, there is a clear need for local community members to engage with existing data and utilize data for their community development (Yoon, Copleand, & McNally, 2018). A few studies give attention to the levels of community engagement by examining community-based organizations’ (CBOs) need for rich data sets with social indicators—such as health, socioeconomic well-being, geospatial, dwelling conditions and safety, and homelessness prevalence data—given the context of CBOs focus on communities’ economic, service, and capacity development (Johnson, 2015; Thinyane, Goldkind, & Lam, 2018). Despite the need for such data and its potential to contribute significantly to communities’ economic and social development, studies identify many barriers to utilizing data due to communities’ limited access to it and data infrastructure, inadequate skills and expertise, lack of leadership, and issues with data quality (Federal Reserve Bank of San Francisco & the Urban Institute, 2014; Mohammed, 2015). Further, smaller communities’ frequent exclusion from the current data ecosystem is because of existing technological, technical, cognitive, and practical barriers (Bhargava et al., 2015).

Considering the significance of data for informing community development, and the lack of existing research examining data’s potential at the local level, this study aims to investigate the process of communities’ data use for community development and decision-making—focusing on the opportunities and challenges of CBOs. We have three specific research questions:

1. How does data support community development and culture?
2. What concerns do individuals have regarding the use of CBOs of data about their communities by CBOs?
3. What challenges do CBOs face during their data use process?

Our approach to community is place-based. A large body of scientific evidence indicates that where people live matters for their well-being (Ellen & Turner, 1997), and, thus, they become sensitive to the changes, development, and success of their geographically defined area. We particularly focus on CBOs as a way to examine data use at the local level. While conceptualizations of CBOs can vary, we modified the definition from the National Community-Based Organization Network (n.d.) to define broadly CBOs as organizations characterized by locally-defined needs and services from locally-based and directed program designs, implementations, and evaluations. The findings of this study will address the challenges and opportunities for developing community-inclusive data ecosystems. As interest in data-driven decision-making grows, achieving an understanding of data potential, workflow processes, and local-level challenges will be necessary for building a culture of data inclusion, which will enable greater inclusion of more communities, especially traditionally underserved communities in data initiatives. While our study is U.S.-centric, our findings could also draw implications relevant in international contexts, given that interests and efforts toward data use for community development is worldwide.
LITERATURE REVIEW

Open data are generally characterized as easily accessible, machine readable, free or available at a negligible cost, and open-licensed (Bhargava et al., 2015; Mohammed, 2015). However, there are efforts to narrow this characterization. Some add a condition, stating that open data should be nonprivacy-restricted and nonconfidential data provided by public money (Janssen et al., 2012); others underscore anonymized data as being open data (Open Data Dialogue, 2012), and still others emphasize interoperability (Davies & Perini, 2016). Open Definition (n.d.) is a bit broader in its assertion that open data is that which “can be freely used, modified, and shared by anyone for any purpose” and “subject, at most, to requirements that preserve provenance and openness.” Despite the difference in approaches, the consistent idea is that open data has made data freely usable to motivate scholarly, economic, and community innovation. While many open data initiatives have been driven by open government movements and data from government and public sectors, open data also includes research and scientific data (particularly those from public-funded research) and private sector data that have both public and private value; for example, vehicle tracking data (Mohammed, 2015).

Many studies discuss the various benefits and opportunities of open data. For instance, open data can increase the transparency and efficiency of governments and businesses, which radically improves their accountability (United Nations Development Programme, n.d.). Access to public data encourages civic participation, helps monitor corruption, and facilitates making demands by tracking the quality and progress of services (United Nations Development Programme, n.d.; Jofre et al., 2016). Further, these data are important components of economic and social development, since they form the basis of policy and administrative decisions influencing social and economic welfare (Castro, 2014; Heeks, 2000). Several studies underscore the potential of data, at local levels, in this regard and illustrate successful cases of data use by community organizations (Gurstein, 2011; Kassen, 2013). Bertot et al. (2014) mention existing national scale data sets from the U.S. Census Bureau (USCB), the Bureau of Labor Statistics (BLS), the Environmental Protection Agency (EPA), and the Centers for Disease Control (CDC), which include varying levels of local granularity—with more localized information at neighborhood, city, and county levels—and can be used in many local community contexts. Thakuriah et al. (2017) argues, particularly from the example of business transactions, for the significance of open data and opinion-monitoring systems—such as real estate, food, house energy, water consumption, or customer reviews—to understand a community’s patterns. The recent development of civic technology further accentuated open data’s potential for various types of community development, which enables citizen engagement with data and makes government more accessible, effective, and efficient for the economic and social good of society (Bhargava et al., 2015). Kingsley, Coulton, and Pettit (2014) discuss the impact of reducing the cost of data assembly, storage, manipulation, and display, which enables many local governments to build their internal data capacity.

Despite the promises of open data and their potential to society, particularly in smaller, local communities, studies also argue that there is a discrepancy between this potential and the added value for stakeholders. One major barrier widely discussed by previous studies is the issue of limited access. Not all data are publicly available or systematically shared with everyone (Jofre et al., 2016). The proprietary nature of data formats is still a problem for data accessibility, especially when formats owned by corporations have wide usage (Jofre et al., 2016). Further, not all stakeholders of open data—such as citizens and minority populations—have equal access to data because of the different digital infrastructures and financial or educational resources that enable data access (Gurstein, 2011; Heeks, 2000). Even if data are publicly available through public repositories, Lafia, Turner, and Kuhn (2018) argue that data are often siloed and difficult to discover. The government provides the data but does not have the resources or the charge to make that data usable. Therefore, third-party entities are or could be relied on, rightly or wrongly, to make the data accessible (Robinson, Yu, Zeller, & Felten, 2008). Johnson, Sieber, Scassa, Stephens, and Robinson (2017) explore the costs associated with making publicly available data usable. Among the (indirect) costs identified, relevant to this study, are the uneven provision of data across geographies and user types and subsidizing private sector economies that in turn create opportunities for corporate influence on government.

Another major issue is the lack of data literacy among citizens, especially from underserved populations or resource-poor communities (Jofre et al., 2016; Magalhaes, Roseira, & Strover, 2013; Sawicki & Craig, 1996). Generally understood as the set of abilities around the use of data (Wolff, Gooch, Montaner, Rashid, & Kortuem, 2016), data literacy includes the abilities to understand, find, collect, interpret, visualize, and support an argument using data, while ethically using the data (Prado & Marzial, 2013). Veeckman et al. (2017) point out that the lack of such abilities and skills makes citizens rely mostly on intermediaries (e.g., journalists, researchers, and so
on) to select and interpret data on their behalf. Yoon et al. (2018) highlight the role of data intermediary organizations in supporting citizens’ data work; however, data literacy education is still critical so that data may be used for citizens’ own benefits, because cases have been reported of data being used against underserved communities, which may only increase social divides (Gurstein, 2011). Bhargava et al. (2015) further articulate the non-inclusive nature of existing data ecosystems, foregrounding data literacy as a critical component for using data to create actionable knowledge, which also supports citizens’ abilities to navigate local information ecosystems.

While a growing body of literature addresses different aspects of open data potentials, benefits, and limitations, studies indicate the lack of systematic, structured, and empirical research in this area (Janssen et al., 2012). While several pieces of research report the empirical uses of open data (e.g., Hausenblas, 2009; Napoli & Karaganis, 2010), these studies tend to regard large-scale, urban areas (Janssen et al., 2012; Martin et al., 2013; Thakuriah et al., 2017) and national scales in the context of developing countries (Chiliswa & Mutuku, 2015; Contreras, 2012; Lee et al., 2011; Mody, Murray, Dooley, & Hospenthal, 2006; Van Schalkwyk, Canares, Chattapdhyay, & Andrason, 2016). Only a few studies have examined local-level, community organizations or marginalized communities (Bopp, Harmon, & Voids, 2017; Johnson, 2015; Kassen, 2013; Thinyane et al., 2018; Yoon et al., 2018). Smaller, local communities can be more data-driven, but the demand and pressure for them to be data-driven, without any necessary support to include them in the existing data ecosystem, canresultingly disempowering them, leaving them only a burden (Bopp et al., 2017). Additionally, local data, given its unique context, is likely not scalable or interoperable with other larger data sets, which in and of itself is not necessarily a bad thing. However, these characteristics could make it harder to attract funders to support local data use. To fully empower communities for data benefits, it is critical to understand the communities’ processes of working with data, the challenges they face, and how existing data ecosystems can support their data use.

3 | RESEARCH METHOD

This study takes a grounded theory approach, using interpretive and phenomenological methods to explore CBOs’ data use processes, challenges, and opportunities. An interview-based, grounded theory approach is well suited to our study, as we aim to understand the practices, experiences, and implications of data in organizations whose work focuses upon community development. We also acknowledge that the meanings, beliefs, thoughts, and characteristics of the specific phenomena we investigate find root within the particular context of the study participants’ viewpoints (Leininger, 1985). Although our study focuses on context-specific meanings, rather than generalizing findings to a population, our approach also offers opportunities for application in other settings (Lincoln & Guba, 1985; Willis, 2007).

To find community organizations that utilized data for their community developmental work, we first identified data intermediary organizations in different U.S. cities, being that data intermediaries are known to support communities’ data use and work closely with community organizations (Yoon et al., 2018). We used a Google search and a review of the 42 partner organizations across 30 cities included in the National Neighborhood Indicators Partnership (NNIP, http://www.neighborhoodindicators.org/), which is a collaborative effort between the Urban Institute (http://www.urban.org/) and local partners to further the development and use of neighborhood information systems in local policymaking and community building. From these results, we compiled a list of 92 organizations and assessed the fit of each organization for this study. After careful evaluation, we worked with nine data intermediary organizations in small (New Haven, CT; Providence, RI), medium (Columbus, OH; Indianapolis, IN; Milwaukee, WI), and large cities (Boston, MA; Dallas, TX; San Antonio, TX). We first conducted a brief interview with the staff in these data intermediaries to understand the nature of their work in supporting communities’ data work and asked for them to connect us with their community clients or partners. From their recommendations, we interviewed a total of 25 participants from different types of CBOs that were analyzed for this study.

Following a predeveloped interview protocol designed to address the study’s research problem, we conducted a semistructured phone interview with these 25 participants. We explored the participants’ experiences working with data using specific example projects, the challenges they faced, the local assistance they received in doing the projects, and so on. While we encouraged participants to discuss their most recent data projects, they also talked about any other projects when they deemed it necessary.

All interviews were audiorecorded and fully transcribed by professionals. The researchers assigned each transcription a control number. The data were inductively analyzed to develop a set of codes, using interpretive qualitative content analysis methods (Hefferon & Gil-Rodriguez, 2011; Mayring, 2014). This exploratory analysis aimed to provide new insights into a social
phenomenon that has not yet been clearly studied. The unit of analysis was each data project, as more than one participant could discuss the same project. High-level codes used in the analysis include: organizational missions, role of data for community work, types of data used, data reuse process, challenge, and local data support. We also developed subcodes, as we saw themes emerging. All analysis used the qualitative data analysis tool, NVivo (QSR International, Burlington, MA). The intercoder reliability among the two coders was 93%.

3.1 Organizational Profile, Participants’ Demographics, and Data Sources

We interviewed a total of 25 participants from various types of CBOs for this study. All the CBOs were non-profit, with missions related to different components of community development, including education, health, economic growth, safety, literacy development, domestic violence, culture, and the environment. The types of CBOs also varied, including community foundations, impact initiatives, educational organizations, religious organizations, public libraries, and community service organizations. We interviewed mostly upper-level administrators, such as CEOs, directors (executive or associate), and, in a few cases, we interviewed program administrators (managers or program directors) or community liaisons when the organizations’ upper administrators referred them to us. Table 1 displays the organizations’ characteristics, particularly regarding their size, budget, and programming focus.

Among our participants, 16 were female and 9 were male. Most of them were in their 40s and 50s (69%), but we had one or two participants from all other age ranges between 20s and 70s. Most of them had extensive experience in community organizations and in-depth knowledge of their own organizations, although their years at the organizations varied greatly, from 1 year to 16 years. The average number of years at the specific organization was 7, but the average number of years in the community development field was 15.

Participants listed different data sources at different levels (e.g., county, state, and national) that they utilized in their community work. While there was a great variety of data sources mentioned, some of most frequently mentioned data sets included: U.S. Census, American Community Survey, economic data from the Bureau of Economic Analysis, data from the Department of Education, KIDS COUNT data, BLS, data from the Department of Transportation, and Department of Health data. Many also utilized local level data, such as the city school district data, regional voter records, regional Children’s Services data, and community health data. While most data sets were quantitative, several also used Geographic Information System (GIS) and experimental data sets.

4 RESULTS

4.1 RQ1: How Does Data Support CBOs Mission to Nourish Community Development and Culture?

All the CBOs in this study have a dedicated mission to nourishing community culture by improving community employment, health and wellness, education, social services, and neighborhood revitalization. Several of them specifically emphasized their role in supporting vulnerable populations—for example, racial minorities, refugees, and those affected by poverty and violence. All participants agreed that the “use of data is central to [CBOs’] mission[s] to understand what’s really happening” (C15), particularly with regard to (a) diagnosis of community problems and community service designs, (b) securing funding and verifying needs, (c) program evaluation,
(d) highlighting community strengths, and (e) fostering policy changes.

### 4.1.1 Diagnosis of unseen community problems and community service designs

Data provide evidence for diagnosing community issues, as the data “show the true stories of what’s happening in the community” (C16). Data also help correct known facts or perceptions and reveal hidden problems. For instance, C08 showed an example of an Asian women’s health issues tendency caused by ignorance of the health sector in his community. He said: “people who work in public health are unaware of the issue concerning Asian women, and [it is] not given [an] equal amount of attention and funding,” despite the fact that “Asian women [have] the lowest use rate for preventative care … like mammogram[s], pap smear test[s], mental health counseling or colonoscopy[es].” He further articulated the reason behind this situation, saying that it is because people usually have incorrect perceptions of, or believe stereotypes about, the Asian community and use data at superficial levels. “Asian Americans have[ve] the highest insurance coverage rate” in his region, despite the ethnic differences among East Asian, Southeast Asian, and South Asian ethnicities, as well as language and cultural barriers, which were “the key factors that prevent Asian women from accessing care and support.” C06 echoed this sentiment, stating that examining existing data through a community lens clarified and opened up new areas of investigation when she tried to identify data contract workers’ employment rates.

Diagnosis of community problems leads to appropriate community service designs. C25 said that after reviewing communities’ literacy rates, high school graduate rates, and employment data, as well as industry sectors’ growth potential, her organization was able to identify adults’ existing levels of basic education, which then led the organization to create a workforce development program. Not only community-level data, but also national-level data from the Department of Labor, were used to determine whether that job training program aligned well with national trends. C10 affirmed this, stating that “data [are] really used to determine” where his organization will ultimately reach out further.

### 4.1.2 Securing funding and verifying needs

Our participants included funders and grantees, and both critically used data in the CBOs’ grant activities. Many CBOs in our study are small-scale nonprofits, which are always in need of financial support. Often, applying for a grant from community foundations is a primary resource for their work. Many said that data provided evidence to support their service plans, to help them be more competitive, and to justify their needs. Several foundations equally underscored the importance of data in grant applications—specifically for verifying applicants’ financial needs. C03 understood the importance of grants in supporting CBOs’ work and would like to “evaluate grant requests and … [make] good grant-making decision[s] by better understanding the community.” C21 argued that verifying data to understand community needs was a critical part of the grant review to determine if “what we’re funding is having an impact.”

### 4.1.3 Program evaluation

Program evaluations and outcome measures are important to demonstrating the CBOs value to their communities and ensuring their accountability to their communities and funders. While CBOs do some data collection on their own to track their outcomes, several also said they consider “statewide and city-wide data to look at impact” and “compare that to the data we collect in-house on people to see if we actually improve people’s lives in [any way]” (C07).

### 4.1.4 Highlighting community strengths

While many participants in this study overwhelmingly discussed the role of data in identifying and addressing communities’ weaknesses, only one mentioned data’s role in highlighting communities’ strengths. C20 pointed out the problem of data being traditionally used to focus on weaknesses and discussed the need to present the communities’ strengths, saying:

> While things could be bad[,] there are always strengths … in a community, and in neighborhoods and in [cities], but we don’t necessarily see that when data is collected, and analyzed, and reported. We tend to see all of the negative indicators, and we never see [any] of the strengths.

C20 further argued for the importance of finding community strengths using data and sharing them with the community “…to get people to want to be engaged and do something about the [community] problems that they know are prevalent in their neighborhood” because it
builds “a sense of connectedness, and a sense of shared responsibility at the neighborhood level … and meeting some of those basic needs means … you can get them engaged and excited about doing [community work].”

4.1.5 | Fostering policy change

A few participants shared their experiences working with policy-makers to urge policy change. C25 said “basically what we do [is] use data to inform policymakers[,] so they can make wise decisions….. [W]e and many people value having concrete numbers when forming policy rather than just relying on anecdotes.” While the pre-school data C25 used may already be available to policymakers, he argued for the importance of adding community context and bringing in community perspectives.

You always hear stories about legislators tend[ing] to be older white men[,] and we don’t have a very diverse group of people who are in charge of the legislature or who get elected to office. So[,] we don’t want to just have people depending on people they run into at the grocery store or articles in the newspaper that appear, which are stories they find extremely valuable. [But] data alone usually does not motivate people to do anything differently…. Having really accurate information about how many people this is relevant for or applies to [in our community] that we can share with the policymakers helps them understand the scale of the program…. It helps when people are doing budgets, which [is] a fair amount of time. (C25)

4.2 | RQ2: What Are the Concerns of Individuals Regarding the Use of Data About Their Communities by CBOs?

Despite all participants explaining the significance of data in supporting their community work and organizational missions, several also expressed skepticism around the idea of open data and using data for community work. C19 said some people were “generally very skittish about the idea of open data” because data are not the raw information but are, rather, contextual. Looking at data through communities’ lenses is critical for accurately interpreting data, and community members felt that the data needed to be accompanied with stories that were alive and supported the data. Because, as C19 said, “[right now] we’re hyper-focused on data … we need to take a step back and start bringing the data alive.”

A few also noted the idea of public trust in data. C12 said, 

There [are] a lot of trust issues around the whole town…. One of the people who was speaking during the public comment was saying how she didn’t really care about the data and what the data said, and I was just … cringing inside.

One potential reason behind this woman’s statement on not caring about the data is that data collected about and from the community are not always shared with community members. When community members participate in the data collection and even ask for it to be shared, they do not always receive a report, which especially frustrates them. C12 considered having access to data as a right of community residence and had concerns that continuous denial to data access would influence the general community’s attitude toward data.

I’ve asked for what data is going to come back to me. I’d like to learn about the air quality that you just tested in my backyard. Or I’d like to learn what you’re finding at the child study centers. Not necessarily about my kid but just in general like what you’re learning as you [conduct] all these studies…. Basically they’re like, “Yeah, we don’t really give the data back.”…. As a resident of [name of town], that’s not acceptable. I can definitely see how the community, after years and years of dealing with that mentality, [is] kind of done with it. (C12)

Others also echoed the issue of public trust around data use. C13, concerned about data being politically used, potentially against the community, said, “we were presented with the issue that community partners don’t see it as an unmitigated positive to have data dissemination, because it can be so politicized. In other words, if a journalist had access to the raw data, how might they use it? If a political opponent had, and so on and so forth, you know?”

In addition, a few participants mentioned the data bias that may influence data collection under certain political paradigms.

A lot of times data do not include undocumented people in a lot of ways. According to the city data, Puerto Ricans are the second- or
third-largest Latino population in the city. The truth of the matter is, they’re really not. Undocumented people who now have to confront the government, they’ll say that they’re Puerto Rican, because Puerto Ricans are natural citizens. They feel more safe being able to say that, because nobody’s going to check on that.... There are very few places that have specific data that we can really use. A lot of it has to do with federal policy, which we don’t work on. (C07)

For C16, “the struggle is really just knowing that we’re working in the paradigm where we can’t collect the gold standard of data.” Further, data bias may stem from the difficulties of collecting certain types of data or using the right measurement because, often, “what is feasible” is the basis for data collection, although it should “not just pick what’s easy to gather but really what’s going to show impact.”

4.3 RQ3: What Challenges Do CBOs Face During Their Data Use Process?

Along with the general concerns around open data for community work, participants also discussed various challenges they faced during the actual process of utilizing data for that work. Four broad themes of challenges emerged from our analysis: (a) disparate approaches to issues of data access, (b) limitations of existing data in the community work context, (c) communities’ widely varying data capacity, and (d) the exclusive nature of the existing infrastructure for communities’ data work.

4.3.1 Disparate approaches to data access

Most participants reported difficulties in data access. While some data mentioned in this study were publicly available online with free access and did not require any further assistance, other data, even though they were “open,” required some communication from agencies (data producers or repositories). C04 observes that “open data” is not really being open because “[the agency] has a database that is not on any of their websites ... [but] it is publicly available.” Knowing that the agency had the data, C04 acquired use of it but noted that “the general public would have no way of knowing that the dataset even exists.” The staff capacity, in such agencies, also becomes an issue, as several participants reported time gaps in obtaining responses from agencies regarding their data requests. C25 said “the hardest thing [was] they have like a million data requests[,] and they don’t have enough staff[,] and they have to prioritize” when working with government agencies. C24 reported a similar experience, which generated one-and-a-half-year gaps in their research team’s data analysis, recalling it as “a bad set of circumstances.”

Data confidentiality and privacy concerns also limit data access. While those data cannot be entirely open, CBOs’ work inherently involves using some confidential and private data, such as “domestic violence and sexual assault, the confidentiality pieces” (C17). In general, the process of accessing data with confidential and private information was very difficult “because of the HIPPA and FERPA and just the general nervousness of the lawyers and the state agencies[,]... and each state agency has a totally different role” (C25). The participants understood the difficulties in general, which were, necessarily, to protect individuals, but the process around data sharing and use agreements could be much improved. C11 argued for a systematic approach to ease the process of accessing confidential data—instead of individual consultation, which never guarantees that the researcher connects to the right authority:

Why can’t we be a little bit more sophisticated and just set up systems and authorities and security levels to say ... “Yes, you can come and see this portion of our data[,] and we’ll set it up[,] and we’ll adhere to the use.” You know, contracts ... and agreements.

C25 further articulated the complexity of data sharing agreements:

[I]f you get the Department of Education to sign a data sharing agreement [that] does not mean the Department of Health is just going to go along with it. They have to start all over again with their own lawyer[s] and their data sharing agreement may be different than the Department of Education[‘s].

Participants also discussed issues of interoperability. Data files, software, formats, organizations, and the use of indicators or measurements are not consistent across data. These inconsistencies not only hinder data access—especially for data software or file formats that are not widely used—but they are also cumbersome and time-consuming for checking feasibility. C21 said, “When you’re an organization that has a small capacity as it is, with big expectations put on you, I think sometimes that gets pushed to the back burner.”
4.3.2 | Limitations of existing data in the community work context

Most participants also discussed the limitations of working with existing data. One of the most common issues participants reported was the need for micro-level data that really reflected neighborhoods and localities. Many argued in favor of “block level data rather than a census track level” (C21), which can demonstrate statistical significance in a smaller area. For example, C03 pointed out that census data is not very community-friendly because the most micro-level data available in the census are at zip-code-levels, which are “not particularly helpful in wanting to understand what’s really going on in a neighborhood,” and “their estimate methodology just doesn’t uncover how quickly a neighborhood might be changing.” C11 echoed this, saying: “many of the public datasets have limitations in what you can do in using [them] ... to get down into neighborhoods as much as possible or low levels of geography.” Often, public data sets also aggregate a certain population into too large a group, and community needs are “masked by the amount of minority [mixing],” which makes “the existing data that [the] U.S. Government has ... misleading” (C08).

In addition, sometimes data are outdated. C07 said “government data tends to be, which most data [are] anyway, tends to be already old by the time we get it, because the government, census data is ten years.” Although C07 understood “it’s hard to measure who’s left the city,” when studying gentrification, time gaps that are too large do not provide useful information, and, as C25 confirmed, it is always challenging to get “the exact data points we want.”

Doing community work inevitably requires CBOs to work with some private data. C04 particularly noted the usefulness of working with private data, such as community members’ “GIS maps of their storm water systems that include information about the depths the pipe is in, what the size of the pipe is, what the material is, etc.” However, using private data reveals two important issues. First, “data sets are not very trustworthy [as they] frequently have errors [because] they don’t have a regulatory context.” Second, although C04 thought these types of data were and should be public, private data must be used through “exclusive sharing agreement[s because of the] security concerns.” C04 continued, “we have a silent protest against it ... we think that’s a smoke screen for reducing oversight and public attention.”

4.3.3 | Communities’ widely varying data capacity

A number of participants realized that “more data [are becoming] public, like through the city and county, [and] there’s been a push toward open data” (C20). However, they also acknowledged simply having “access to data doesn’t mean that you know actually what you’re looking for” (C20), noting CBOs’ limited capacity in working with data. A few mentioned the challenge of time commitment when doing data work, considering that their organizations are small and have limited human resources available for data work because of the time required for daily operations.

A more prominent issue with community capacity has to do with an organization’s data literacy. All participants were aware of the importance of skills and expertise for using data: “If you don’t have knowledge or expertise in what you’re actually searching for or how to do a query, you can get the wrong data without knowing that that’s not actually the data you need.” (C21)

General data literacy levels also varied among the CBOs in this study. While one participant stated, “we can’t understand what the number means” (C08), others said they were developing skills and were now able to understand but did “not always know how to explain it to [others]” (C11) because they were “generalists” (C15). While a few organizations in the study have an expert staff member or team of data experts, more than one-third said they did not have the necessary data skillset and struggled with data interpretation. C13 stated,

Certainly, most community organizations don’t have the [bandwidth], the time to work with data at that level, and [many] are very frank in saying they don’t have the skills. They weren’t trained as researchers, and there’s nobody on their staff who is either. So[,] even if they had this raw data, what could they possibly do with it in house[?]

Participants saw data literacy as “a special skillset,” which “takes a lot of training and specialization” (C17). Often, they felt that nonprofit community organizations do not have enough financial resources to either hire data experts or train existing staff because “it’s really expensive to have the skillset on staff and pay for it” (C17), and they “can’t necessarily afford the [cost]” (C21).

Perhaps for these reasons, several participants noted that they had to or preferred to rely on others when lacking expertise and necessary skillsets in their organizations. C21 said,

[In our opinion[, working with a data intermediary is so helpful because that intermediary should be able to help you define what, in the landscape of all the data that are]
available, what type of data would be most helpful in what you’re trying to measure.

4.3.4 | Exclusive nature of existing infrastructure for communities’ data work

Several participants explained how existing infrastructures are not very supportive for communities using data for community work. C09 felt that many data owners generally “don’t invest in the creation of an infrastructure” that supports data sharing and (re)use for communities. She further said, “if they do, typically they want to own it[,] they have what I jocularly call a Sauron strategy, one ring to rule them all, they want a big[,] heavy weight[,] frightening[,] centralized data repository,” which is often unrealistic. She criticized most existing data exchange systems as “still crippled on the backend … and there’s often some budget hungry proprietary software firm gatekeeping access to the data model of how the whole thing actually works.”

A few others shared experiences of getting no response when requesting that existing data be made more community-friendly. When C25 tried to “correspond some demographic information to the data that we put in” and tried to “drill down a little deeper on the state level,” the response was always negative, and C25 did not have any idea if the state agency was unwilling to help or did not have the skills.

A small number of participants who had experiences working with academic partners reported some difficulties due to different interests, as communities were more interested in short-term, practical solutions, while academic researchers were more geared to longer-term, theoretical work. Often, a lack of data literacy skills created a communication barrier with academic researchers. C07 said “[professors] have referred me to books[,] or they’ve given me a quick 101 kind of things” to help her understand basic data literacy, but “I still don’t feel comfortable to [be confident].”

In addition, C12 pointed out the inequality in partnership and resource distribution when conducting community research that limited her organization’s capacity-building in the data realm. She said:

We just got asked to partner with three very small community non-profits here in [our neighborhood] on a USDA grant, and it’s not a huge grant, about $100,000 a year… [O]nly $4,000 … goes to our indirect, so that leaves very little money to actually offer any staff support or other resources to get the evaluation…. I think that’s a major, major institutional problem[, and] I intend to try and figure out if there’s a solution to that. I’ve been really thinking about what partners I could reach out to on the national level to figure out how they handle that.”

5 | DISCUSSION

Our findings empirically support existing literature that discusses the benefits and potential of data at local levels. Data play an important role in supporting community organizations’ missions in different ways, particularly in diagnosing community problems and designing community services, securing funding and verifying needs, evaluating programs, highlighting community strengths, and fostering policy changes relevant to community development. As previous studies indicate the inherent difficulties in measuring the changes and impact of data use at local levels (Kassen, 2013; Yoon & Copeland, 2019), presenting the empirical use of data for community

| TABLE 2 | Open data challenges in community context with opportunities |
|---|---|---|---|
| **Level** | **Challenge** | **Community suggestion** | **Opportunity** |
| Infrastructure | Data access; exclusive nature of existing data ecosystems | Create community friendly and inclusive data infrastructure | Create partnership among local data ecosystems |
| Social | Public trust around open data | Invite community participation; treat community as an active player in data ecosystem; include community voices | Community engagement; collaboration with local governments |
| Community | Lack of local context in existing data; community (organizational) capacity | Community-driven local data collection; build community capacity on data | Community engagement; community research; educational programming |
| Individual (community member) | Data literacy; individuals’ data capacity | Data literacy training | Educational programming |
development can be one way of demonstrating open data’s usefulness and further encouraging more data use at local levels, while planning for long-term measurements of its impact on communities.

Along with these findings concerning benefits and usefulness, our research also identified four broad themes of challenges: disparate approaches to data access issues, limitations of existing data in the community work context, very few people with the data skills necessary to effectively use data for community decision-making exist in any given community, and the exclusive nature of existing infrastructures for communities’ data work. The challenges identified in our study demonstrated that the barriers to community’s data use were not at one level; rather, they are multidimensional at different levels, which requires different approaches at the individual, community, social, and infrastructure levels. Understanding communities’ challenges at those levels also helps to identify opportunities for the improvement of collaborative work with communities and their stakeholders. Table 2 presents identified challenges associated at varying levels with opportunities.

The infrastructure level challenges have the most challenges than any other level. These challenges mostly stem from the exclusive nature of existing data ecosystems for communities’ data use, since the current data ecosystem does not always offer a doorway to understanding, interpreting, and managing data-driven decision-making for communities. Additionally, it does not incentivize or empower communities to control their own data and its uses (Bhargava et al., 2015). Community members also reported that existing data access systems are not always community-friendly, as they often experience difficulties in acquiring data in a timely manner and speaking with data owners for proper help. Existing practices and policies around data sharing and reuse sometimes hinder a smooth workflow for data access, especially when data are protected or proprietary. While most existing, open, public data are government data, community members felt that more data should be considered public—especially when that data has public value. However, changing perceptions concerning data with public and community value, from private to open-public data and providing easier access to those data, still has a long way to go.

At the social level, community members showed great concern about open data, due to the lack of public trust around data collection, representation, and use. In general, they felt that communities were often not invited to the open data movement, especially when data collection was designed to address community issues or when data represented community. Even if they were part of a community data collection and representation, they were often seen as a passive player and, thus, excluded from local data collection design; instead, they rather passively participated in a given setting, without a chance to ask questions or raise their voices to reflect their opinions. The tensions and controversies generated by this passive participation from the top-down approach in community engagement are not new (e.g., Stuedahl, Runardotter, & Mörtberg, 2016; Yoon et al., 2018) and made the communities question the openness and deemed desire and trust in open data context.

At the community level, many participants also indicated that existing data are limited in supporting communities’ needs. The level of granularity and locality are not always reflected in existing data sets, and the current data supply chain does not support the often time-sensitive nature of community work. Perhaps this is why many community organizations face needs and shortcomings in their own data collections. Previous studies have also reported that community organizations, with or without partnership with academic or data intermediary organizations, collect local data to complement these limitations in existing data sets, but not all are equipped with the proper financial or human resource capacities (Yoon et al., 2018).

Both at the community and individual level, capacity building around data use and data literacy skills appeared to be a major challenge. Despite the importance of data literacy skills in using data for community developmental work and decision-making, many CBOs in this study discussed the difficulties of training staff or hiring data experts within their organizations. Many were also unaware of existing, local resources for staff education or training, which may be partially due to the lack of such opportunities in their local communities. Several noted their preferences for hiring consultants, which they felt to be a more cost-effective choice, but this was not an option for all CBOs in this study. It was not that the CBOs lacked interest in developing data literacy skills, but, rather, many CBOs lacked financial and human resources. They could not sacrifice existing staff members’ time for training, nor did they have the budget for such training. Opportunities for local training—offered for free or at reasonable costs—seemed to be their preference.

Because our sample was selected through local data intermediaries’ networks, many CBOs in this study had, at least, knowledge of or access to data intermediaries and expressed the usefulness of having them locally and of working with them. However, while local data intermediaries’ roles are significant in supporting communities’ data work (Kingsley, Pettit, & Hendey, 2013; Pettit, Hendey, Losoya, & Kingsley, 2014), not all communities have data intermediaries with the same levels of
While our study presents some significant challenges of communities’ data use, several opportunities also emerged to address each challenge, as several participants in our study also suggested. Those opportunities include building partnerships for a community-friendly data ecosystem, promoting a culture for community engagement and community-driven research, and meeting the educational programming needs of the community. At the very bottom level, creating educational programs that meet communities’ needs seem to be the most urgent agenda, given the challenges that CBOs in this study experience. It is also critical to support the community’s capacity building for data use. Different sectors developed and provided such programs in the last decade. The coverage of those programs varied, such as online data literacy education for individuals (e.g., Data Equity for Main Street project, https://data-equity.org/), hands-on workshops to learn data literacy concepts at libraries (e.g., plix’s DataBasic.io, https://plix.media.mit.edu/activities/databasic), data culture building at an organization for their capacity building (e.g., Data Culture Project, https://databasic.io/en/culture; FabRiders, https://www.fabriders.net/), organizations receiving training to help the staff leverage data and technology to tackle local priorities (e.g., Data 101 course, https://www.neighborhoodindicators.org/data-tech/course-catalog/data-101-data-visualization-data-literacy-and-storytelling), and provisions for social learning around data for individuals and organizations’ data readiness (Data Playbook, https://www.preparecenter.org/toolkit/data-playbook-toolkit).

Despite the growth of educational opportunities around data, many community members in our study noted that they were unaware of such opportunities at the local, national, or international level, which suggests the need to promote existing opportunities, especially the ones that are freely available. Further, Bhargava et al. (2015) argue against the un-inclusive nature of the data literacy concept itself, as the existing concept focuses on particular skillsets in a certain technology platform. This un-inclusive nature suggests the need to revisit the concept and redefine it in order truly to meet the needs of communities when developing any training and educational programs—especially those of underserved communities. By equipping them to understand the underlying principles and challenges of data, they can empower communities to comprehend, interpret, and use their own data that they encounter or produce.

Engaging communities more with an existing open data movement will also address the numerous challenges and concerns that many of the study participants experienced at social and community levels. Community organizations in this study expressed a strong desire to be included in the open data movement and play more of an active role. By participating in passive data collection within a given setting, as well as proposing a new data collection needed to address community issues, these organizations play the role of a watchdog in other stakeholders’ data use regarding their community. It is not surprising that several participants expressed community concerns around data when the existing data ecosystems were not very community-friendly. For example, participants indicated that existing data are not localized sufficiently to truly reflect their communities’ natures, while they strongly felt that adding community context to data interpretation was critical. Further, communities’ fears concerning data—that is, whether data will be collected without any personal or political bias, will be used for the communities but not against them, and will be transparently shared with the communities—caused an issue of trust that should be properly addressed early on, so that open data initiatives may move forward.

The involvement of community representatives in existing data ecosystems, as well as bringing their voices and concerns into the process of data collection, distribution, and use/reuse, seems to be one way of addressing this trust issue. While it is not always clear how often community members receive invitations to existing open data initiatives, or if they have been only seen as end users, solidification of trust will only be achieved through collaboration and equal partnership among different stakeholders in the data supply chain—such as data collectors, aggregators, distributors, intermediaries, policy-makers, data designers, users, and end users. A co-created or bottom-up approach to open data initiatives, by implementing shared decision-making and equal partnership, will be necessary.

Finally, our study suggests that it is critical to extend or build partnerships among various stakeholders for supporting communities’ data literacy education and beyond. It further suggests the importance to provide a strong support to a community’s data work and to create a more community-friendly data ecosystem at a local level. Higher education, local knowledge organizations, such as public libraries, local data intermediaries, local governments, as well as other local entities all have roles to play. Higher education can greatly contribute to data literacy education, given the higher education field’s teaching expertise as well as its organizational mission to serve community and civic engagement. While some of the study participants expressed existing partnership models due to the unfair resource distribution and different nature in their approaches toward community work,
successful models also exist that demonstrate the potential role of higher education to support community’s data use. A number of higher education institutions developed close relationships with communities through community engagement research, which can be easily transformed or implemented according to data focus. Further, working with higher education practitioners, who specialize in research design, is an opportunity for communities to produce data that they often want to and need to collect in order to supplement existing open data (Yoon et al., 2018). Public libraries, using their experience and expertise in teaching information literacy skills, also have great potential to provide training programs for communities to learn how to identify and evaluate relevant data for the communities’ own use and analysis. Data Navigators 2.0 at the Providence Public Library is a good example. While data intermediaries have a long history of serving communities for development and data needs, recently, libraries and data intermediaries have sought partnerships between themselves (e.g., see Civic Switchboard: https://civic-switchboard.github.io/), which is a good model for developing partnerships within existing, local data ecosystems. The Data Equity (http://data-equity.org/), a joint project of the California State Library and the Washington State Office of Privacy, who partnered with San Jose State University, is another example of how different stakeholders in existing data ecosystems, such as governments, libraries, and higher education institutions, can work together to support communities’ data literacy and beyond. Although these projects are recent, they serve as solid examples, showing the synergy of team efforts by building reliable and sustainable models for long-term partnerships and impact on community work.

Based on our findings, we have listed recommended best practices for each stakeholder in community data ecosystems. We acknowledge that there could be other types of stakeholders of community data ecosystems, depending on the context of communities’ data work, but the list below would be the starting point to build the model for community inclusive data ecosystem engaging stakeholders that may be involved in any community’s data work.

- Data provider/collector/local government
  - Apply best practices of data curation to make data more reusable.
  - Work with other stakeholders to provide equitable data access to community members.
  - Be aware of community needs and community data users.
  - Co-design local data collection with community members.
- Data repository
  - Work with data providers to make data more easily accessible to community members.
  - Be mindful of various types of data users such as underresourced communities when developing data services.
- Data intermediary
  - Support community data use by providing data services.
  - Support community’s data literacy training/education.
  - Support local data collection that meets the community’s need.
- Higher education
  - Conduct community research through equal partnership with community.
  - Provide data literacy training/education for community.
- Public library
  - Provide programming/education around data and data literacy.
  - Provide access to local data collection to their serving community.
  - Support local data collection that meets the community’s need.
- Community data user (e.g., CBO/citizen)
  - Build capacity in data through partnerships and collaboration.
  - Be an active player in community data ecosystem.

6 | CONCLUSION

The purpose of this study was to investigate the process of communities’ data use for community development and decision-making—focusing on the opportunities and challenges that CBOs face during their data work for communities. Our study presented data’s role in supporting communities’ development by: diagnosing community problems and designing community services, securing funding and verifying needs, evaluating programs, highlighting community strengths, and fostering policy changes relevant to community development. Despite data’s role, our findings suggested that several major challenges hinder CBOs data use: difficulty accessing data, limitations of open data (un-local nature, excluding essential data from being open), limited data capacity (especially in data literacy skills), and difficulties using and accessing existing data infrastructures. These challenges suggest the un-inclusive nature of current data ecosystems for communities. However, considering that all open data initiatives are still recent, there are opportunities for addressing these challenges, particularly by
creating educational programming, building partnerships within data ecosystems, and bringing community voices forward in current data ecosystems. Addressing these challenges cannot be a short-term task, but it is better to consider them early to keep open data initiatives moving forward. Also, by addressing these challenges, CBOs can truly realize data’s potential for all citizens by strengthening communities and improving individuals’ lives.

ACKNOWLEDGMENTS
This project was funded by the Institute of Museum and Library Services (#LG-96-17-0184-17). The authors also thank the nine data intermediaries and NIPP for their help with data collection and Paula McNally for her help with initial data analysis.

ORCID
Ayong Yoon https://orcid.org/0000-0002-3312-1234

REFERENCES
Aronson, B. (2004). Improving online access to medical information for low-income countries. *New England Journal of Medicine*, 350(10), 966–968. https://doi.org/10.1056/NEJMfp048009
Attard, J., Orlandi, F., Scerri, S., & Auer, S. (2015). A systematic review of open government data initiatives. *Government Information Quarterly*, 32(4), 399–418.
Bentley, C. M., & Chib, A. (2016). The impact of open development initiatives in lower- and middle-income countries: A review of the literature. *The Electronic Journal of Information Systems in Developing Countries*, 7(1), 1–20. https://doi.org/10.1002/j.1681-4835.2016.tb00540.x
Bertot, J. C., Jaeger, P. T., & Grimes, J. M. (2010). Using ICTs to create a culture of transparency: E-government and social media as openness and anti-corruption tools for societies. *Government Information Quarterly*, 27(3), 264–271. https://doi.org/10.1016/j.giq.2010.03.001
Bertot, J.C., Butler, B.S., & Travis, D.M. (2014). Local big data: The role of libraries in building community data infrastructures. In *Proceedings of the 15th Annual International Conference on Digital Government Research - Dg.o ’14* (pp. 17–23). New York, NY: Association for Computing Machinery.
Bhargava, R., Deahl, E., Letouzé, E., Noonan, A., Sangokoya, D., & Shoup, N. (2015). Beyond data literacy: Reinventing community engagement and empowerment in the age of data. Retrieved from http://datapopolance.org/item/beyond-data-literacy-reinventing-community-engagement-and-empowerment-in-the-age-of-data/
Bopp, C., Harmon, E., & Voida, A. (2017). Disempowered by data: Nonprofits, social enterprises, and the consequences of data-driven work. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI ’17* (pp. 3608–3619). New York, NY: Association for Computing Machinery.
Castro, D. (2014). The rise of data poverty in America. Retrieved from http://www2.datainnovation.org/2014-data-poverty.pdf
Chiliswa, Z., & Mutuku, L. (2015, May). Building open data infrastructure and strategies for effective citizen engagement. Paper presented at the 2015 Open Data Research Symposium, Ottawa, Canada. Retrieved from http://www.opendataresearch.org/dl/symposium2015/odrs2015-paper13.pdf
Contreras, J. (2012). Open access scientific publishing and the developing world. *St Antony’s International Review*, 8(1), 43–69.
Davies, T., & Perini, F. (2016). Researching the emerging impacts of open data: Revisiting the ODDC conceptual framework. *The Journal of Community Informatics*, 12(2). Retrieved from http://www.ci-journal.net/index.php/ciej/article/view/1281
Ellen, I. G., & Turner, M. A. (1997). Does neighborhood matter? Assessing recent evidence. *Housing Policy Debate*, 8(4), 833–866. https://doi.org/10.1080/10511482.1997.9521280
Federal Reserve Bank of San Francisco and the Urban Institute. (2014, December 4). Data use is key to improving communities: New book challenges policymakers, funders, and those seeking to broaden opportunity to work more strategically. Retrieved from https://www.frbsf.org/community-development/blog/data-use-improving-communities-new-book-challenges-policymakers-funders-opportunity-work-strategically/
Gurstein, M. B. (2011). Open data: Empowering the empowered or effective data use for everyone? First Monday, 16(2). Retrieved from http://firstmonday.org/ojs/index.php/fm/article/view/3316
Hausenblas, M. (2009). Exploiting linked data to build web applications. *IEEE Internet Computing*, 13(4), 68–73. https://doi.org/10.1109/MIC.2009.79
Heeks, R. (2000). *Government data: Understanding the barriers to citizen access and use* (Working Paper 10). Retrieved from https://www.researchgate.net/profile/Richard_Heeks/publication/334638131_Government_Data_Understanding_the_BarrIers_to_Citizen_Access_and_USE/links/5d374ee24585153e591c4063/Government-Data-Understanding-the-Barriers-to-Citizen-Access-and-Use
Hefferson, K., & Gil-Rodriguez, E. (2011). Interpretative phenomenological analysis. *The Psychologist*, 24(10), 756–759.
Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. *Information Systems Management*, 29(4), 258–268. https://doi.org/10.1080/10580530.2012.716740
Jofre, A., Szigeti, S., & Diamond, S. (2016). Citizen engagement through tangible data representation. *Foro de Educación*, 14(20), 305–325. https://doi.org/10.14516/fde.2016.014.020.015
Johnson, M. (2015). Data, analytics and community-based organizations: Transforming data to decisions for community development. *U.S. A Journal of Law and Policy for the Information Society*, 11(1), 49–96.
Johnson, P. A., Sieber, R., Scassa, T., Stephens, M., & Robinson, P. (2017). The cost(s) of geospatial open data. *Transactions in GIS*, 21(3), 434–445.
Kalampokis, E., Tambouris, E., & Tarabanis, K. (2011). A classification scheme for open government data: Towards linking decentralised data. *International Journal of Web Engineering and Technology*, 6(3), 266–285. https://doi.org/10.1504/IJWET.2011.040725
Kassen, M. (2013). A promising phenomenon of open data: A case study of the Chicago open data project. *Government Information Quarterly*, 30(4), 508–513. https://doi.org/10.1016/j.giq.2013.05.012
Kingsley, G. T., Coulton, C. J., & Pettit, K. L. S. (2014). *Strengthening communities with neighborhood data*. Washington, DC: Urban Institute.
driven society. *The Journal of Community Informatics, 12* (3), 9–26.

Yoon, A., & Copeland, A. (2019). Understanding social impact of data on local communities. *Aslib Journal of Information Management, 71*(4), 558–567. https://doi.org/10.1108/AJIM-12-2018-0310

Yoon, A., Copleand, A., & McNally, P. (2018). *Empowering communities with data: Role of data intermediaries for communities’ data utilization.* Retrieved from https://scholarworks.iupui.edu/handle/1805/17971

**How to cite this article:** Yoon A, Copeland A. Toward community-inclusive data ecosystems: Challenges and opportunities of open data for community-based organizations. *J Assoc Inf Sci Technol.* 2020;71:1439–1454. https://doi.org/10.1002/asi.24346