Chapter 15
Urban Governance

Alex D. Singleton and Seth E. Spielman

Abstract In this chapter, we discuss how the availability of new urban data has the potential to transform the governance of cities. Such effects are realized in several ways: by increasing transparency; creating greater scope to appropriately set and measure municipal policy outcomes; and by design of well-planned and managed digital infrastructure, better empower citizens to hold decision-makers to account. However, such potential is not without risks, and without critical reflection, the proliferation of new data and their integration into software delivering algorithmic insight or automation may reproduce or develop new inequalities. We conclude that for digital urban governance to make a future that we want, it is important that we reflect upon how and where these technologies are implemented to ensure these are optimized in favor of the public good.

Keywords Governance · Dashboards · Open Data · Algorithms

15.1 Transparency and City Open Data

Transparency in the processes of city governance both limit the potential for corruption, while also ensuring that the citizens of urban areas can hold democratically elected officials to account for their use of public funding. UN Habitat (2004) argues that greater transparency can reduce urban poverty and enhance civic engagement; and by promoting engagement through a range of different policy instruments, can reduce citizen apathy, make service delivery better contribute to poverty reduction, increase ethical standards, and grow city revenues. Transparency within urban governance is an expansive topic. However, we focus here on the role of Open Data within this context.

A. D. Singleton (✉)
Department of Geography and Planning, University of Liverpool, Liverpool, UK
e-mail: alex.singleton@liverpool.ac.uk

S. E. Spielman
Department of Geography, University of Colorado Boulder, Boulder, USA
e-mail: seth.spielman@colorado.edu

Data about our cities are legion and include both traditional sources such as surveys or censuses, and those new forms of data related to other collection mechanisms such as sensors (e.g., noise, pollutants, etc.), social media, or as an operational by-product (e.g., meeting minutes, expenses, administrative records). The ownership and control of access to such data are a key facet of transparency, and much data about cities are held within the private realm. For example, geolocated Tweets posted by citizens of urban areas is owned by the private company Twitter, with public access restricted to either limited subsets of Tweets or commercially procured full access. The costs of accessing these data may, however, be prohibitively expensive to all but a few users. By contrast, Open Data are distributed under very different licensing conditions, typically enabling data to be supplied without cost, and to be reused and re-distributed without downstream licensing implications. Within some countries, an Open Data license has a more formal definition; for example, the UK adopts an Open Government License (https://www.nationalarchives.gov.uk/doc/open-government-license/version/3/) for officially defined Open Data.

There are several common rationales given for the release of Open Data. The first is to provide a resource that can enhance civic engagement in the processes of governance. For example, through the provision of data about the expenses of government employees, these are open to scrutiny and oversight. Secondly, Open Data can be integrated into platforms design to improve aspects of public service (e.g., school and healthcare comparison). Finally, Open Data can act as a driver for innovation and has the potential to create both direct and indirect economic benefits. Despite such diverse potential benefits, the release of Open Data is however not free, as the preparation, maintenance, and hosting of data assets have costs attributed (Spielman and Singleton 2015; Johnson et al. 2017). Furthermore, their release or availability is often governed by complex political data economies. For example, the permanence of Open Data can be somewhat illusionary, and there are examples of where Open-Data licenses have been revoked retrospectively and for future releases, or where guidance associated with such a license has been adapted so that this constrains future use. In the USA, the removal of the website open.whitehouse.gov followed the election of Donald Trump; and in the UK, the Land Registry switched its policies for data previously distributed with an Open Government License to terms that are more restricted.

15.1.1   Open Data Platforms

Within many municipalities, Open Data are disseminated through online portals, with two popular platforms including Socrata (https://www.tylertech.com/products/socrata) and CKAN (https://ckan.org/). An example of an Open Data platform running CKAN is shown in Fig. 15.1.

There are a number of reasons why such data portals provide better tools for transparency over simply sharing data through a static Web site. Most platforms provide access to search, highlighting the breadth of the available data; and results
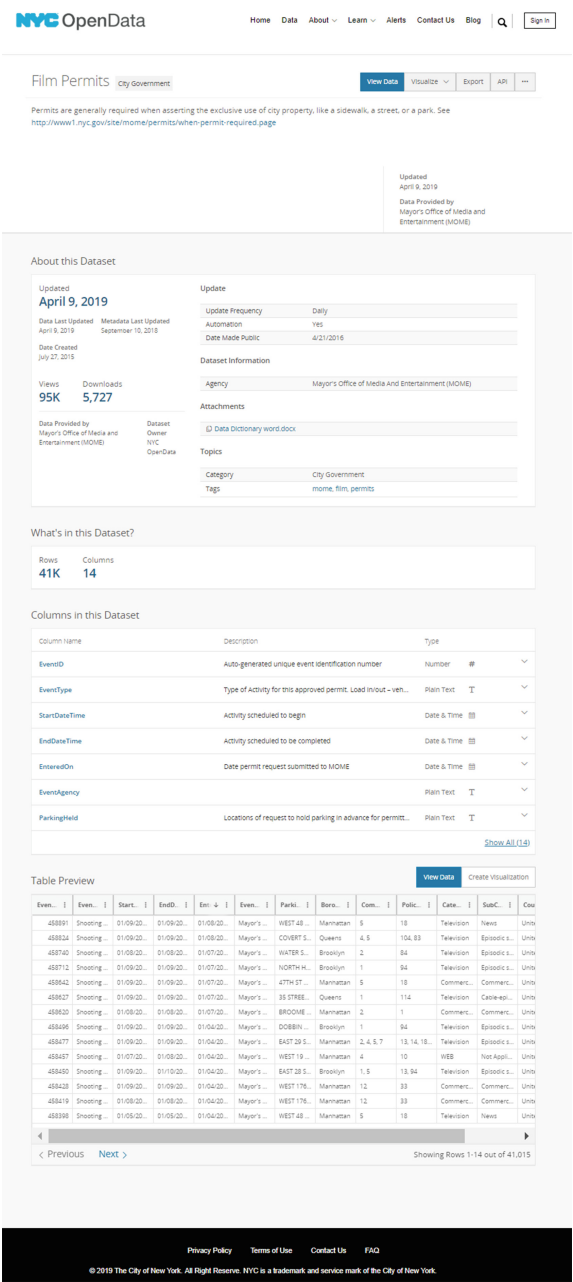
Fig. 15.1 Open Data portal for New York City showing a catalog entry for film permits

are typically returned alongside detailed metadata, sample extracts and some limited visualization capability. With many portals, data sit within a database that, in addition to being presented to the catalog's visual interface, are often also made available through publicly accessible application programming interfaces (API), enabling integration into a wide variety of software and tools. Such API endpoints and associated document object identifiers (DOIs) provide permanent and direct links to Open Data that enhance both usability and reproducibility.

However, the extent to which a community can benefit from engagement with sources of Open Data or those platforms designed to turn these assets into information can be variable; and differences may manifest between social, racial, ethnic, and economic groups. Mitigating access differentiation has to be a priority in urban governance if the implementation of Open-Data systems is to be maximized in the interests of the public good.

However, it is important to recognize that the creation of effective Open Data platforms requires significant investment. Organizationally, it is complex to initiate buy-in from stakeholder data owners, and additionally to facilitate the creation of effective management, storage, dissemination, outreach, and training associated with such new data infrastructure investments. Glasgow, which is the largest city in Scotland, was the recipient of £24 m of government funding to deliver a Future Cities demonstrator project (Sarf 2015). Around £7 m of this investment was allocated to build "Open Glasgow," which is a data platform providing access to numerous and previously siloed urban data. The project involved making 372 different datasets available through a CKAN-based Open Data portal alongside an online mapping platform provided by Esri. Around 21 different roles were associated with this project, and beyond the technical implementation, included additional support for Open Data development, engagement, and hackathons.

15.1.2 Open Data and Accountability

The growing adoption of Open Data platforms is a positive development, but in and of themselves these platforms have little impact on the lives of citizens. To have an impact, Open Data platforms have to be used by people and organizations. This means that the usability and accessibility of the platform itself are essential, but more importantly, it means within either the city agencies or the public at large, there must be constituencies who have the skills and time to transform the data assets into information.

The potential benefit of Open Data is only realized if certain conditions are met. We argue that Open Data repositories for urban governance should follow a set of principles that are accepted by scientific communities. These are sometimes referred to as the FAIR principles: findable, accessible, interoperable, and reusable.

- Findable: Data are published to stable and publicly accessible URLs. The URL is advertised and made known within the government and across agencies.

- Accessible: Data should be published in a usable format with stable and well-documented procedures for access. For example, pdfs are not a usable format for data. Access protocols should be well-documented and standardized; for example, APIs should remain stable over time. Individual data files should have a static URL. Data have to be documented and documentation must be maintained.
- Interoperable: Data should be organized such that linkages between data sets, and/or over time, are possible.
- Reusable: Data should have licensing provisions that allow flexible reuse of data.

In an effort to boost engagement some data-savvy communities sponsor events to encourage public consumption of the data published on open platforms. A consortium in New York City regularly organizes events around Open Data. For example, in Boulder, CO, USA, the city sponsored an "Art of Data" exhibition which encouraged local artists to create physical works of art from digital data. Some forms of digital data, such as text, can be difficult to work with in traditional forms of analysis—in the City of Boulder's Art of Data Exhibit, one artist built an installation based on individuals' test responses to survey questions about safety and other aspects of city life. Creative use of public data can be strikingly impactful. However, getting residents or the public, private, and not-for-profit sectors to use Open Data, and to communicate their findings to a broader audience, can be difficult yet is critical to closing the loop and allowing Open Data platforms to achieve their potential. Incentivizing creative use of data seems like a wonderful way to spur innovation; however, the lack of well-established norms of use and goals for Open Data platforms inhibits the impact of these resources.

We believe that the most impactful uses of public data focus on accountability; that is, using data to track progress toward institutional, individual, or collectively defined goals. However, there are not well-established models around how Open Data platforms might be integrated with participatory social and political processes to guide and track progress at the city-scale. Identifying and tracking progress toward goals can be non-trivial in the urban context.

Cities are large and complex systems bureaucratically, physically, and socially. Developing an understanding of the components and their interrelationships within systems is enormously difficult. For the average citizen, it can be hard to know where a city's responsibilities begin or end and observing the scope of a city's operations in a particular domain can be very difficult. Cities are a patchwork of public and private land, with city agencies often having overlapping jurisdictions and conflicting priorities. For example, a transportation department might want to increase the number of vehicles moving through an intersection and the planning department might want to improve pedestrian safety by reducing traffic volume. Given such organizational complexity, assessing accountability and progress toward goals can be complex. Goals may not be shared between various parts of the city's administrative structure. Moreover, the institutional goals may not be shared by the residents of the city, and in some communities, residents may have different priorities than others.

Open Data potentially simplifies some of this complexity by providing citizens and other interested groups with mechanisms to observe these large systems and to

understand where cities are, and are not, investing resources. That is, if the right data are made available at the right level of aggregation, citizens can begin to observe the city not just as the space within which their daily activities take place but as an organizational unit.

Here we focus on the conditions required to realize the potential for Open Data to improve governance and in particular to drive accountability; in this context, using data systems to track progress toward measurable social and organizational goals. While ideally, these goals would emerge from participatory public processes, we omit discussion of these mechanisms here.

15.1.3   Why Are Goals Important?

Simply stated, the concept of accountability as applied to public data is that citizens (and municipal leaders) can hold public-sector agencies accountable for their work. However, large and complex projects that are undertaken without clear goals can be difficult to assess. For example, consider the partnership between Kansas City, Google, Sprint, and Cisco to develop a highly instrumented corridor with WiFi and advanced traffic control systems. In spite of millions of dollars in investment, it is difficult to say whether the project has been successful. The media report that the project reduced travel time an average of 37 s. Sprint, as a company, harvested data from thousands of citizens. But did the project achieve its goals? Was it a success? If so, for whom? Without clearly stated and measurable criteria, it is difficult to answer such questions.

A framework of accountability can, however, have powerful and positive social impacts. When police departments around the USA started to publish data about the racial characteristics of people they stop and question, glaring social inequalities were laid bare. In cities across the USA, data highlighted and confirmed the long-running perception that racial minorities in the USA are disproportionately targeted by the police. The use of Open Data to hold police departments accountable for seemingly biased patterns of enforcement is an excellent example of citizen empowerment in the challenge of existing doctrines. Our implicit goals in this example refer to widely held beliefs around how public institutions ought to function; for example, that enforcement of laws should be uniformly applied, not based on race or class.

15.1.4   Dashboards and Performance Indicators

Open Data dashboards simply make data or information available to municipal stake-holders. Data in their raw form are only consumable by people with those technical skills (and time) to both effectively frame questions and then investigate. Dashboard interfaces provide a more widely accessible visual interface to data. Often, a dash-board will display indicators that are derived from data. An indicator can be simple

and direct, such as the number of traffic citations written in the preceding 30 days, or complex and derived such as the social vulnerability of the population. Kitchin et al. (2015) document the spread of the dashboard and its increasingly widespread use around the world. They critically argued that rather than simply "reflecting cities, [dashboards] actively frame and produce them." Whether they are mirrors reflecting data or instruments of power seems secondary to the fact that dashboards are widely used, and in governance, they can be used productively or unproductively.

In and of themselves, dashboards accomplish very little. They find their utility through linkage with implicit or explicit social goals and incorporation into some governmental process that links action (or incentives) to the indicators on the dashboard. A dashboard that simply displayed data, disconnected from meaningful administrative or social goals, would have little impact. For example, to provide insight into racial bias, the police department in Minneapolis, Minnesota, USA, publishes a dashboard breaking down police stops by race, location, gender, and age (https://www.ins idempd.com/datadashboard/); while this dashboard is not linked to explicit goals and targets, it is squarely addressing implicit social goals. On the other end of the spectrum, the City of Boulder, Colorado, USA, uses a dashboard to track progress toward explicitly stated targets around safety, health, livability, sustainability, housing, and governance (Fig. 15.2). While rudimentary, the dashboard uses a simple system of green checks for targets that are met and red exclamation points for missed goals. A public process determined the indicators to be tracked on the dashboard; these were derived from the city's "Sustainability and Resilience Framework" which was a document designed to guide "budgeting and planning processes by providing consistent goals necessary to achieve Boulder's vision of a great community and the actions required to achieve them" (https://www-static.bouldercolorado.gov/docs/Sustainab ility_+_Resilience_Framework-1-201811061047.pdf).
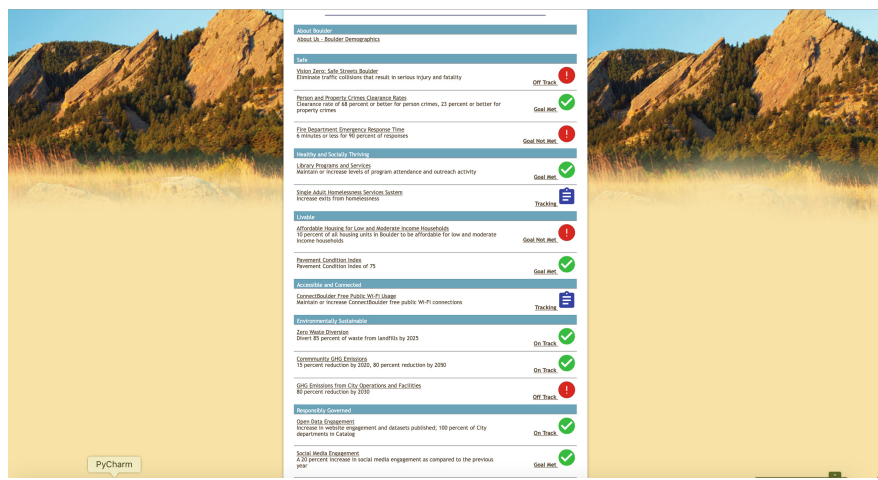

Fig. 15.2   A goal-based dashboard from the City of Boulder, Colorado, USA

The use of quantitative targets, such as those employed by Boulder, is a widespread practice in the private sector where such indicators are sometimes called key performance indicators (KPIs). Performance indicators are powerful tools in so far as several criteria are met:

- Urban KPIs must measure the right things. That is, they must quantify social, political, or economic processes of interest to the leaders and residents of the city.
- Urban KPIs must be actionable; measuring things that residents and leaders have no power to change is of no consequence. Dashboards should in some meaningful way drive action.
- Urban KPIs must be correctly measured: Data quality is a serious concern for public dashboards. Linking data to public goals creates incentives to manipulate or misreport data.
- Not all goals are quantifiable: It is important that KPIs and dashboards play an appropriate role. Critical social goals, such as well-being, may be unmeasurable but this does not mean that public institutions should not strive toward them.

There are, however, critiques of dashboards and urban data more broadly, notwithstanding that it seems to us that they are rooted in a genuine effort to provide transparency and accountability. While data may be imperfect and the social processes that produce them may be loaded and flawed, we strongly argue that providing access to information is better than not. Dashboards, when made public, reflect a kind of self-imposed, publicly stated accountability toward targets. While it is true that measuring what matters to the residents of a city is a non-trivial exercise, and that data systems are more likely to reflect things that can be measured than things of direct concern to residents, there is some meaningful overlap. It is within this space of overlap where data can help advance the governance of cities.

15.2 Algorithmic Decision-Making

There is a proliferation of increasingly granular measures or insights that can be extracted from urban data, which is necessitating new methods for both their management and their analysis. Algorithms are computational processes that are designed to solve a particular problem, which within an urban context can relate to both aspects of urban analytics (e.g., which communities are best served by green space), or the implementation of operational models (e.g., traffic light control systems). Algorithms can also have differing degrees of autonomy through their specification, estimation, or implementation. The use of computational algorithms within urban contexts is not new, and they have a lengthy history of application, from models applied to make predictions about the spatial organization of human activities, to those teasing out geodemographic structure from multidimensional spatial data (Webber 1975), alongside those which have been implemented operationally to guide decision-making (Foot 1982).

15.2.1 Positioning Algorithms

The argument is made that the successful implementation of algorithms can augment or supplant human expertise. For example, a fire inspector may have knowledge of the city in which he or she works and might choose buildings to inspect based on his or her expertise. Alternatively, an algorithm might rank buildings based upon the probability that they contain a building code violation. In one realization of an algorithmic process, an inspector could be dispatched to all buildings scored as risky by the algorithm. Alternatively, the algorithm could augment the inspector's expertise, providing him or her with a way to guide attention. In either case, the use of algorithms in law enforcement raises questions about the biases, fairness, and transparency in algorithms, especially when algorithms are trained or validated based on historically biased enforcement actions.

We believe that there are three broad use cases for models and algorithms in urban governance. By models, we mean tools that use learned or estimated parameters to produce classifications, probabilities, or scores. Algorithms are computational procedures that may or may not involve data and models. We use the two terms somewhat interchangeably, preferring the term "algorithmic decision-making" to refer to the use of computation to augment municipal operations. The use cases for algorithmic decision-making are:

- Augmentation: This refers to the use of models to guide or enhance human expertise. For example, using machine learning to augment the building inspector's expertise and to help focus efforts on buildings likely to contain a violation.
- Replacement: Using an algorithm in place of a human: for example, using combinations of cameras and radar to automate traffic enforcement. In this case, the machines determine if a violation occurred and take action. The computational enforcement system replaces a human system.
- Efficiency: Using models or algorithms to manage urban systems. Computation enables a kind of dynamic optimization that is difficult in the absence of sophisticated systems. For example, heating, ventilation, and air conditioning systems in buildings may take into account occupancy, outdoor temperatures, historical norms, and other factors. Transport systems may make small adjustments to signal timing system-wide in order to continuously adapt to variations in traffic and demand, thus optimizing flow.

At their best, across these use cases, algorithms potentially present an unbiased way to improve public welfare and the operation of cities. That is, well-designed systems can make people safer and urban systems more efficient. Machines potentially remove individual biases and capacities from urban management and enforcement. When algorithms and models are transparent and interpretable by humans, they move decisions out of the subjective and political domain into the public sphere. Open algorithms and models can also force conversations about principles, such as what kinds of actions or places should be targeted, or what publicly generated training or validation data should be used. Such models can then embed these collectively

generated principles. Enforcement actions are then the result of a public process around the kinds of factors that contribute to risk or that the community wants to minimize or maximize.

At their worst, algorithms could become super enforcers of institutional biases and racism, and reinforce existing structural inequalities, or at the extreme create new ones. When algorithms replace humans (or are positioned at the extremes of augmentation), there are valid concerns that the system-automated surveillance that emerges violates basic human rights to privacy and equal (unbiased) enforcement of laws. For example, it is not possible to place surveillance cameras everywhere; from the perspective of a police department, placing cameras in high-crime areas might be an efficient use of limited resources. However, if algorithmic tools are used to augment enforcement or replace policing it means that people in high-crime areas have a higher probability of being found guilty of crimes than those in areas without cameras, even if algorithms are fair and unbiased.

15.2.2 Challenges for Operationalizing Algorithms

Unlike inferential models that have historically been applied within urban contexts, many contemporary and emerging methods from the cannon of data science, AI, and machine learning focus instead on prediction, which produces models with operational utility, but because the structural manifestations of causal effects are often hidden, their value can be argued as limited in terms of explaining how processes operate over time and space, and as such, we have weaker understanding of the dynamics of systems. Although we may be able to make very good forecasts from such new modeling paradigms, this is in tension with generalizable models of how the world functions, and the development of theory.

Additionally, many new algorithms that are used to create predictions rely on big data that are used to train models, which is the process by which an algorithm learns from the past to make new or future predictions. However, in doing so, an analyst has to be certain that there are no systematic biases in such data, and that any measures taken are likely to be stable over time. The non-compliance of such issues has been argued as integral to cases where previously successful models stop making effective predictions: for example, inaccuracies in magnitudes predicted by Google Flu Trends (Lazer et al. 2014).

Beyond issues of measurement, it has also been noted that most if not all big data are socially constructed, which also leads to potential bias, and should drive ethical considerations and framing. If such data are integral to the function of algorithms, and those decisions that they advise or take, the algorithms themselves can inherit such same bias; and as such may ensue real-world implications if adopted uncritically (Kitchin 2014). For example, the content of social-media data is only representative of those people who generate it, and so may under- or over-represent certain socioeconomic or demographic characteristics; or for georeferenced data, accuracy may be impacted by both where the social-media data were collected (e.g., the built

environment impacting GPS signal reflection) or by people's prevailing attitudes to location sharing. More generally, crowdsourcing refers to the process of the public contributing attributes of observed phenomena for some particular purpose. Such data collection does not have an a-priori sample design, and as such the data's underlying collection is influenced by those who engage with a project. For example, the Street Bump (https://www.streetbump.org/) application was created for the city of Boston, USA, and collected data using the accelerometer in phones when a depression in a car was recorded as it passed over a pothole. These readings were pooled and analyzed to identify where remedial action may be required on a street. The representativeness of such data was, however, bound up in the collection process, with the application only being available to those with an iPhone, those who could afford one of these handsets, and additionally a subsection of this population who would be likely to install the application, and additionally volunteer geolocated information. Such a segment of the population may also have particular travel patterns, and there is additionally potential that only a partial survey of the city is conducted through such a tool. Understanding such bias and how this might impact algorithmic governance is a fundamental issue that should be considered by decision-makers.

15.3   Conclusion

In this chapter, we have outlined how the processes and operationalization of urban governance are being enhanced and challenged through the emergence of new digital technologies that relate to the instrumentation of cities, how those data being generated, and how the information derived can be used within urban contexts to enhance decision-making. For digital urban governance to be effective we posit that the inclusion of stakeholders by design, aligned to principles of transparency and openness, is essential in order to mitigate risks of associated negative dystopian consequences. The power of new digital frameworks has great potential to improve the health, prosperity, inclusivity, and sustainability of cities; yet it is essential that these technologies do not end up reinforcing past injustices, or at their most extreme create new inequalities. Future cities will be digitally augmented, and the challenge for us now is to critically reflect on the impacts that ensue from these new technologies, and to make sure we plan for a future that we want.

References

Foot D (1982) Operational urban models. Taylor and Francis, London

Johnson PA, Sieber R, Scassa T, Stephens M, Robinson P (2017) The cost(s) of geospatial open data. Trans GIS 21(3):434–445

Kitchin R (2014) The real-time city? Big data smart Urbanism. GeoJournal 79(1):1–14

Kitchin R, Lauriault TP, McArdle G (2015) Knowing and governing cities through urban indicators, city benchmarking and real-time dashboards. Reg Stud Reg Sci 2(1):6–28

Lazer D, Kennedy R, King G, Vespignani A (2014) The parable of Google Flu: traps in big data analysis. Science 343(6176):1203–1205

Sarf N (2015) Learning from future cities demonstrators. https://innovateuk.blog.gov.uk/2015/10/29/learning-from-future-cities-demonstrators/. Accessed 22 May 2015

Spielman SE, Singleton A (2015) Studying neighborhoods using uncertain data from the American Community Survey: a contextual approach. Annals Am Assoc Geogr 105(5):1003–1025

UN Habitat (2004) Transparency in local governance. United Nations Human Settlements Programme. Nairobi, Kenya. https://unhabitat.org/books/tools-to-support-transparency-in-local-governance/. Accessed 10 April 2019

Webber RJ (1975) Liverpool social area study, 1971 Data. PRAG Technical Paper 14. Centre for Environmental Studies, London

Alex D. Singleton is a Professor of Geographic Information Science at the University of Liverpool. He has been PI or Co-I to over £14m of research income and is Deputy Director of the ESRC Consumer Data Research Centre (CDRC) and Director of the ESRC Data Analytics & Society CDT. His research is concerned with how the complexities of individual behavior, attitudes, and contexts manifest spatially and can be represented and understood through the framework of Geographic Data Science.

Seth E. Spielman is an Associate Professor of Geography and Information Science at the University of Colorado where he also serves as the Assistant Vice Chancellor for Data Science and Strategy. His research focuses on how to create and reason with digital representations of social and economic space.