URBAN GEOGRAPHY
https://doi.org/10.1080/02723638.2021.2019499

The datafication of water infrastructure and its implications for (il)legible water consumers

Fenna I. Hoefsloot[a], Christine Richter[b], Javier Martínez[a] and Karin Pfeffer[a]

[a]Faculty of Geo-information Science and Earth Observation, Department of Urban and Regional Planning and Geo- University of Twente, Enschede, Netherlands; [b]Fraunhofer Center for International Management and Knowledge Economy IMW, Leipzig, Germany

ABSTRACT
Redevelopments of Lima's water infrastructure aim to reduce inequalities in water consumption, connections, and coverage by implementing data technologies and claim to make urban water management more efficient. However, little research has been done on how the city's hydrosocial geography is shaped by the increasing use of data for the supervision and control of its water infrastructure. This article analyzes the datafication of Lima's water infrastructure as the interplay between different legibility-making practices to understand how the use of multiple, interoperable and real-time data sources, shapes the hydrosocial geography of the city as well as the relationship between Lima's main provider of water and sewerage services (SEDAPAL) and urban water consumers across three scales: newly urbanized areas, water sectors, and households. We conclude that, in an already unequal urban landscape, the datafication strategically (re)structures the relationship between SEDAPAL, as a state organization managing the water infrastructure, and Lima's residents.

ARTICLE HISTORY
Received 3 June 2020
Accepted 3 December 2021

KEYWORDS
Water infrastructure; datafication; Lima; legibility making; smart urbanism

1. Introduction

It is said that when the Spanish colonialists first settled in the Incan empire that has come to be known as the Republic of Peru, they arrived in January, the only time of the year in which the coast of Peru is not covered by a thick layer of fog and the rivers are flowing with water (Leonard, 2000). Wanting a port capital, the Spanish founded the "City of Kings" along the Pacific coast and on the banks of the river Rimac, an area that had been inhabited and irrigated by Incan and pre-Incan communities, not knowing that the city would grow out to be the second-largest city of the world built in a desert (Allen et al., 2017; Leonard, 2000). Today, the City of Kings goes by the name of Lima and structurally faces water scarcity. With more than nine million people and 10 mm mean annual rainfall (Ioris, 2016), it is a constant challenge to guarantee the provision of water to all the inhabitants of the Lima Metropolitan area (Fernández-Maldonado, 2008; Miranda Sara et al., 2016). This is perhaps most clearly

CONTACT Fenna I. Hoefsloot f.i.hoefsloot@utwente.nl Faculty of Geoinformation Science and Earth Observation, University of Twente, ITC Building 3-011, Hengelosestraat 99, 7514 AE Enschede, The Netherlands

noted in the geographical and social inequalities in the water consumption and water connections amongst residents and the water coverage in time (Fernández-Maldonado, 2008).

To address this inequality in water distribution, redevelopments of Lima's water infrastructure aim to make the water infrastructure more efficient, reliable, and equitable by implementing "smart" data technologies (SEDAPAL, 2014, 2015). Similar "smart city" narratives have steered cities worldwide to mimic the image of a high-tech, fully connected, data-driven urbanism (Luque-Ayala & Marvin, 2015; McFarlane & Söderström, 2017). Today, urban society, materiality, and economy have become intertwined with coded algorithms (Lindley et al., 2017; Shaw & Graham, 2017). Specifically concerning urban infrastructures, supervisory control and data acquisition (SCADA) systems have been extensively rolled out in cities globally to monitor and control flows in water, traffic, and electricity grids (Kitchin & Dodge, 2017). These developments have spawned the idea that, with sophisticated and reliable digital technologies and data production, it will be possible to make the measurement and monitoring of infrastructures more efficient and equitable and reduce "human insecurity" in the management and governance of urban flows (Kitchin et al., 2015; Luque-Ayala & Marvin, 2016; Taylor & Richter, 2017).

In the case of Lima, Peru, these expectations of "smart city" technologies are no different. Digital data is assumed to provide faster and better monitoring of the fluctuation in the water pressure, detect leakages quicker, and improve the insights about water usage (SEDAPAL, 2015, 2018; World Bank, 2011). Underlying these efforts to implement digital data technologies in the infrastructure is the wish to manage the water that flows through the city more efficiently. The "smartening" of the water infrastructure is embedded in the modernist ideal of the controllable and legible city, although the means to see the city have changed. Whereas previously, governing bodies relied on cadastral maps and paper-based censuses to monitor the city, today, digital data technologies are seen as essential tools to sense urban dynamics and make urban infrastructures legible in near-real-time (Glasmeier & Christopherson, 2015; Offenhuber, 2017). The premise is that by installing meters and implementing a supervisory system, it becomes possible to construct an informative representation of urban reality and "see" the water flows through the data (Birnholtz & Bietz, 2003). These innovations would then help identify breakdowns and non-regulated tapping and improve water distribution amongst the consumers of the water infrastructure (SEDAPAL, 2015).

However, the idea of the "smart city" has extensively been critiqued, with researchers raising concerns about the often technocratic and neoliberal assumptions of the smart city discourse and concerns regarding the introduction of digital technologies in cities (Vanolo, 2014). Specifically, critics have pointed out the influence of large-scale measurements and big data on surveillance (Kitchin & Lauriault, 2018), the protection of privacy (Elwood, 2008; Kitchin & Dodge, 2017), the bias in the representation of people in the data (D'Ignazio & Klein, 2020), and the flow and possession of data (Taylor & Broeders, 2015).

This article contributes to these debates by analyzing the effects of datafication of Lima's water infrastructure as layered legibility-making practices. As the gaze upon the city has become electronically mediated, we discuss how the hydrosocial geography and the relationship between residents and the state is shaped through multiple data sources. Informed by debates on socio-hydrology and hydrosocial research, we adopt hydrosocial geography as referring to the dynamic and mutual constitutive relationship between society and water (Wesselink et al., 2017) and their relationship with and across urban space. Such discussion is specifically important for Lima, where recent digitalization efforts take place in an environment of persisting and especially pronounced hydrological inequality.

We structure the rest of the paper as follows. In section 2, we frame this research within the broader debates on smart urbanism from a sociotechnical perspective and describe the process of datafication as a practice of legibility making from the perspective of Lima's Drinking Water and Sewerage Service, known by its Spanish acronym SEDAPAL. This is followed by the methodology (section 3). In sections 4 and 5, we describe the datafication of Lima's water infrastructure and analyze how the legibility-making practices on three scales – the mapping of settlements, the supervision of the water distribution system, and the metering of household water consumption – create a differentiated hydrosocial geography with distinct categories of water consumers. Finally, in section 6 we conclude with how layered legibility making is dynamic and strategic in shaping the relationship between SEDAPAL and urban water consumers.

2. Conceptual approach

2.1. The datafication of urban infrastructure

Similarly to modernist approaches to urban planning (Rabari & Storper, 2015), mainstream smart city approaches are characterized by a technocratic view on urban space and processes (McFarlane & Söderström, 2017; Vanolo, 2014) and a strong belief in the potential of technology to sense and register complex urban interactions into quantitative data and translate them into objective and valuable knowledge (Van Dijck, 2014).

In the often corporate-led transition toward smart cities, the process of datafication – referred to as the quantification of daily life (Mayer-Schönberger & Cuckier, 2013) – is presented to achieve a fully connected and governable city. Datafication of urban infrastructure can be described as the trend of increasingly using digital data in larger volumes and at increasing detail in infrastructural management (Heeks & Shekhar, 2019; Mayer-Schönberger & Cuckier, 2013).

The change in data production using digital technologies can have profound implications for the organizational structure and governance of our cities and for urban infrastructures. Critical urbanism has extensively discussed how "smart" technologies and data can lead to reductionism in urban planning and technocratic governance (Luque-Ayala & Marvin, 2015; McFarlane & Söderström, 2017; Verrest & Pfeffer, 2018). From a development perspective, there has been a critical evaluation of the goals for becoming "smart," and how the race to cyber cities is driven by capitalism and results in new global inequalities (Gaffney & Robertson, 2016; Taylor & Broeders, 2015). Moreover, feminist and post-colonial data studies have raised questions concerning representation in data and unpacked how data is inherently biased (D'Ignazio & Klein, 2020). Such studies reveal how data, be it big or small, are never pre-analytical or a-political. Instead, data are the product of a specific context and framed to serve a purpose (Kitchin et al., 2015). Policy, legislation, culture, and market forces steer what is worth knowing and who or what is counted in the city's datafied image.

Therefore, the approach we take in this study shifts attention to the relationships and interaction between the data and the water infrastructure in Lima. With such focus, we follow Karasti et al. (2018), Pfeffer (2018), and Bowker and Star (2000), who view infrastructural development as a process of aligning different forms of knowledge, humans, and artifacts; a process that stabilizes into networks that may re-align again later in time. While being inspired by sociotechnical research of urban infrastructure and smart urbanism at a more general level (for example, Alda-Vidal et al., 2018; Amin & Thrift, 2017; Kaika, 2012; Kemerink-Seyoum et al., 2019; Swyngedouw, 2006), analytically, our study draws specifically on Scott's notion of legibility making to describe datafication in Lima's water infrastructure and its effect on the city's hydrosocial geography. In the following section, we explain the concept of legibility making and how its conceptualization is useful to support our investigation.

2.2. Datafication through the lens of legibility making

In his seminal book, "Seeing Like a State", Scott (1999) uses legibility-making to conceptualize the state-imposed processes of abstraction and simplification of real-world messiness to improve its management. Through different practices and techniques, such as the standardization of measurement, the land cadaster, and the national census, the state maps its territory and registers its subjects. Scott (1999) emphasizes how the state employs legibility making practices as a tool for the state to increase governability of people and territory. Legibility, in this sense, increases administrative order and control over its residents and resources and diminishes alternative knowledges and systems of organizations.

Scott (1999) draws on a variety of empirical examples such as land administration, natural resources management, fiscal government, and urban planning to illustrate how large-scale administrative plans have tried, but ultimately failed, to simplify the complexity of the real world into manageable schemes and categories. However, as Scott's research repeatedly shows, this way of "seeing like a state" cannot account for on-the-ground heterogeneity. As an analytical lens, legibility making helps explore how different categories of people emerge and how, in doing so, the relationship between people and the state is defined.

As such, the concept of legibility making makes for a useful tool in our approach to understand the datafication of Lima's water infrastructure. However, it is also necessary to distill how legibility-making practices in the age of "smart cities" differ from those of high-modernist states, as described by Scott (1999). Li (2005), for instance, argues that Scott's focus on the controlling and authoritarian relationship between the state and its citizens has overlooked the network of public and private actors that are working on the project of legibility making. Already in 2005 Li pointed out that we need to consider also non-state actors in the processes of legibility making. As datafication progresses, the role of private industry and its interactions with state actors in legibility making has further increased (Taylor and Broeders (2015). Today, it is a comprehensive network of organizations, individuals, and technologies that collect and maintain data, and in the process, negotiate to categorize and define urban spaces (Richter & Georgiadou, 2016). As Luque-Ayala and Neves Maia (2019) point out, the rationale for making legible depends on the actors' aims. In Rio de Janeiro, for example, the digital mapping of favelas has been led by corporate actors such as Google and Microsoft, with the self-proclaimed aim of economic integration of the neighborhoods and capturing georeferenced data (Luque-Ayala & Neves Maia, 2019). Whereas legibility making of the state aims at increasing governability, corporate-led legibility making is often driven by economic aspirations.

Additionally, the properties of the data that underpin legibility making practices in the smart city have changed compared to legibility making practices of the modernist city described by Scott (1999). Scott argues that the schemes of legibility and statecraft used in high-modernist states were in their aim not very different from the governance methods used by the societies preceding them. However, what drastically changed in the 20th century was the magnitude and range of the legibility making practices. Through technological progress, the high-modernist state had gained the ability to map whole territories and achieve (near) universal registration of people and resources (Scott, 1999). With the more recent emergence of digital data technologies, the state has again gained new legibility-making tools. Kitchin and Lauriault (2018) unpack how the introduction of digital data technologies has increased the volume, velocity, variety (structured and unstructured), exhaustive scope, resolution, and relational nature of the data that is collected about people. What sets apart datafication as legibility making practice in the smart city compared to modernist legibility making practices is that the data is digitally interoperable (favoring standardization and the exchange of data), layered (favoring a fixed layering and ordering of data), potentially more diverse in source and structure, and sometimes collected in real-time (see, Table 1). In the transition toward smart urban governance, we need to consider the interactions between different databases created and curated by various actors involved in the project of legibility making.

The effect of legibility for the state is, as Kalir and van Schendel (2017) write, a double-edged sword. On the one hand, counting and recording can increase the legibility of, and governance over, the territory it administers. On the other hand, the legibility of processes also increases the state's accountability toward the people who are registered (Kalir & van Schendel, 2017), determines resident's access to urban resources (Anand, 2017; Taylor & Richter, 2017), and reconfigures the alignment of people, data, and agency, rather than solving frictions (Pelizza, 2017). Legibility making, thus, involves the conscious selection of what is recorded and not recorded (Kalir & van Schendel, 2017). The state constantly balances the logic of creating legibility and increasing state administration or maintaining illegibility and outsourcing responsibilities (Gandhi, 2017; Truelove, 2018).

Because the dynamics between legibility and illegibility making are context dependent, they may take different courses in states with high levels of informality and illegibility, for instance, in the Global South. As Truelove (2018) points out, not "seeing" is oftentimes not an arbitrary failure to make legible, but the strategic creation of the absence of the state and its services. In these zones of illegibility, alternative service provision strategies emerge, such as formal or informal markets (Wutich et al., 2016) or auto-constructed infrastructures built and managed by citizens (Hoefsloot et al., 2020). Moreover, through selective legibility and illegibility-making practices, a relationship between the state and its residents is created, and citizenship is determined (Gandhi, 2017).

Table 1. Conceptual overview of commonalities and differences between legibility-making in the high-modernist city (Scott, 1999) and legibility-making in the smart city.

		Legibility making in the high-modernist city	Datafication – Legibility making in smart city
Commonalities	Rationale	Increasing administrative order and control over people, resources, and territories.	
	Means	Large scale data collection schemes focused on standardization and simplification.	
	Effects	Reduction of on-the-ground complexity into manageable categories and the deletion of diversity in knowledge. Increase in accountability of the state toward residents.	
Differences	Rationale	Mainly state-led data collection aimed at increasing governability.	Multiple data collectors/collections and diverging rationales.
	Means	Structured data from few sources. Longer-term periods of collection and updates.	Structured and unstructured data from multiple sources, including digital ones. Data is layered and largely interoperable. (Near) real-time collection and updating.
	Effects	Reductionist and selective regarding relationships between the state, society, and data. Results in dichotomy between legible and illegible.	Differentiation on multiple layers leading to degrees of (il)legibility. *(scope of this paper)*

Therefore, in this paper, we analyze the datafication of Lima's water infrastructure as legibility-making in the smart city to understand how the use of multiple data sources, which are layered, interoperable, and in near real-time, shapes the hydrosocial geography of the city as well as the relationship between Lima's main provider of water and sewerage services (SEDAPAL) and urban water consumers.

3. Methodology

We analyze the datafication of the water infrastructure from the viewpoint of Lima's state-owned drinking water and sewerage service, SEDAPAL. This research is based on 6-month fieldwork conducted from 2019 to 2020 in Lima, which combined qualitative research methods in the form of interviews, focus groups, site-visits, and observations with informal conversations. In total, 12 in-depth interviews have been conducted with staff of different operational departments of SEDAPAL (five), civil society actors (three), and key-experts in the field of water management and governance (five). The focus of the interviews was on the daily operation and use of the digital data infrastructure and the water distribution system. In interviews with engineers of SEDAPAL, we "walked through" the digital data technologies that are used to monitor the water flows in the city, providing a glimpse into the infrastructure from SEDAPAL's perspective.

Additionally, we draw on conversations with residents to understand what is seen and what is overlooked by SEDAPAL through the process of datafication in Lima's water infrastructure. Between December 2019 and February 2020, three focus group discussions were conducted with residents from three districts of Lima (San Juan de Lurigancho, Barrios Altos, and Miraflores). Participants were encouraged to discuss the water service provision in their neighborhoods. Each focus group lasted between 75 and 100 minutes. The three areas were chosen as they reflect, at least to some degree, Lima's range in socio-economic living conditions, (in)formality, and geographical location. The participants in the focus groups were recruited from already existing civil society networks in each district. As a result, many, although not all, participants had been previously active in urban development projects or research initiatives and often held a leadership position within their communities.

The site visits included formal and informal areas in Lima and the main water treatment plant, La Atarjea. The data, the text, notes, and transcripts were coded in ATLAS.ti™ based on an initial code list focusing on (i) the different elements in the water infrastructure (actors, digital technology, data, and the water distribution system) and (ii) the rationale for the design and functioning of the water infrastructure. See appendix 1 for the code book. Coding was conducted by the first author, who also collected the data in the field. The code list was adjusted and expanded during the process of analysis as new themes emerged. Based on this process, the three categories of legibility making emerged that we will discuss in the following sections. These are the registration and mapping of the urban expansion area (4.1), the centralized supervision of the water distribution system (4.2), and the metering of household water consumption (4.3). In the final results section (4.4), we analyze the effects of legibility-making on water access.

4. Results: Practices of legibility making

The legibility making of Lima's water infrastructure can be best explained from the perspective of La Atarjea, Lima's main water treatment plant, and the central node in the infrastructural network. There, from the offices on the site, the engineers of SEDAPAL can supervise the water pressure on the pipes at any given time and place in the primary network; and they can monitor the consumption and payments for the water and detect leakages or breakdowns. In the control rooms, they view the water distribution system through the data provided by sensors and meters installed throughout the infrastructure. These data are collected in a web-based geo-information system and presented in maps, charts, and models representing the water infrastructure. The screens, models, and tables they portray are legibility-making practices materialized. They translate the intricate infrastructural network into an image, which can be read at a glance.

This digital data infrastructure has been developed over the past decades through large-scale, bilateral, or multilateral infrastructural development projects (World Bank, 2011, 2018). The emphasis of these developments has been on creating a more efficient water infrastructure and should be seen in the light of the discussion revolving around the privatization of public services in Peru. During the neo-liberal Fujimori governments in the '90s, and driven by the International Monetary Fund (IMF) structural adjustment programs, many of Peru's public services were privatized (Fernández-Maldonado, 2008; Ioris, 2012). As part of this, SEDAPAL was stripped of many of its previous functions related to infrastructure implementation, maintenance, and repair. In its current "lean"

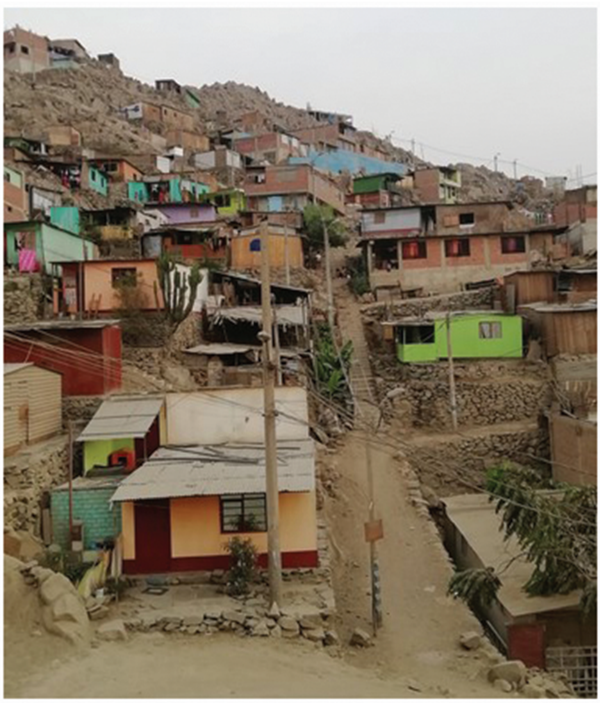

Figure 1. Newly urbanized area in José Carlos Mariátegui in San Juan de Lurigancho, Lima's largest district. Photo was taken by the first author.

form, SEDAPAL has taken up the role of coordinator of the water infrastructure and contracts third parties to execute projects. While the privatization was never completed and SEDAPAL remains under the auspices of the Ministry of Housing, Construction and Sanitation and the Ministry of Economy and Finance, it is still an option discussed publicly (El Comercio, 2019; Ioris, 2012). The main arguments for the privatization of SEDAPAL have centered around the belief that privatizing water management would improve management and give an economic incentive to make the system more efficient. As a response, SEDAPAL has taken up neo-liberal policies and prioritized efficiency for water management (Ioris, 2012).

The process of datafication that we describe in the following sections should be considered within this context as the aims of improving the efficiency of the water infrastructure and reducing non-revenue water have been driving rationales for the legibility-making of Lima's water infrastructure. In sections 4.1, 4.2, and 4.3, we describe in detail legibility-making practices on three scales that amount to the datafication of Lima's water infrastructure.

4.1. Practices of legibility making of unplanned and underserviced settlements

Since 1950, Lima has experienced several phases of rapid urban growth in which people settled on barren lands haphazardly, incrementally forming new city districts in the process (Riofrío, 2003). The lack of data about Lima's unplanned settlements poses a challenge for SEDAPAL in planning for the expansion of the water infrastructure. Therefore, to breach this gap, SEDAPAL uses geo-technology to make newly constructed neighborhoods legible within their cartographic data.

These new neighborhoods (see, Figure 1), referred to by many names in the past (*barriadas, pueblos jóvenes, asentamientos humanos*), range from fairly planned and structured areas, according to the national urban planning regulation, to unplanned informal neighborhoods lacking essential service provision. The residents of these newly established settlements often live in a state of pre-formalization: their presence is acknowledged, albeit not yet registered in the several databases of the municipality and service providers that define formal citizenship. Up to today, the city's newest neighborhoods are generally undermapped (Lambert & Allen, 2016), and service networks such as water infrastructure and electricity often arrive years after the construction of the houses (Criqui, 2015; Fernández-Maldonado, 2008).

This trajectory goes in the opposite direction from the formal developments in Lima, where developers need to receive an expression of intent from SEDAPAL that they will be able and willing to provide water infrastructure to planned developments before receiving a construction license from the municipality (Republica del Perú. Decreto supremo N ° 029-2019-VIVIENDA, 2019). According to the National Institute of Statistics and Information (INEI), 4% of Lima's residents are still pending connection (INEI, 2019). However, due to the lack of formal registration of residents, this number is widely contested. According to local experts, the breach in connection to the infrastructure ranges between 10% to 30% of Lima's real population.

Recently, SEDAPAL started using drones and geo-technology to map the newly urbanized areas and improve the planning for the expansion of the water infrastructure. The aerial images acquired by the drones are digitalized and georeferenced by a tertiary company. In translating these images into a map, one of the main challenges lies in determining the boundaries between parcels. Some boundaries, such as roads and staircases, are visible in the image and can be included in the cartography. Yet, in other cases, it is impossible to demarcate one terrain from the other based on visual features, forcing SEDAPAL's engineers to estimate the boundaries. Each parcel is categorized depending on the type of built-up and the zoning plan of the municipality. This is important since people who have inhabited zones classified as archeological sites or high-risk zones are considered illegal, and SEDAPAL will not be able to provide water infrastructure there. In addition to mapping newly built-up areas, areal images are also used to estimate the terrain which has been prepared for development but where construction has not yet started. This provides an insight as to where future urbanization will occur.

By combining the data from the areal mapping with the maps SEDAPAL received from newly urbanized areas when they apply for a water connection, SEDAPAL is said to have the most up-to-date map of Lima. The maps of the unplanned settlements are meant to aid SEDAPAL in making a more accurate prognosis of the future water demand and speed up the processes of infrastructural delivery. Nevertheless, informal urbanization in Lima continues, and new neighborhoods are constructed on an almost daily basis. Hence, despite speeding up the mapping process by using digitalized areal images, SEDAPAL's cadaster cannot keep up with the reality of urban expansion.

Additionally, while the digital mapping of settlements by SEDAPAL is making new informal neighborhoods legible, in practice, this does not change much for the processing of residents' requests for water provision. In 2004, the "Water for All" (*Agua para todos)* reform made it possible to receive a piped water connected with a "proof of possession" (*constancia de posesión*) rather than a land title. This meant that many people without a land title could receive a water and sewerage connection by handing in proof of possession, accompanied by a perimeter and elevation map approved by the municipality and a neighborhood map, including the exact location where the requested pipes and connections must be installed. The final document must be signed by an engineer of an accredited consulting company and the municipality and handed over in a hard copy to SEDAPAL. As the power to recognize the claim to water by a resident does not lie with SEDAPAL, but with the municipality, it is necessary to follow the formal application route and deliver a paper map. Without the literal stamp of approval and signature by the municipality, the digital map can only inform future plans; it does not serve to start the construction of the water infrastructure in the unplanned settlements.

Thus, although GIS technologies serve as a means to increase the digital legibility of unplanned settlements, due to the lack of legal power of the digital map, these settlements remain illegible in the official cadaster used to manage the water distribution system. The effects of not being legible are significant. It not only makes people underrepresented but also undermines the legitimacy of their claim to resources in the city and their contribution to infrastructure development. For example, SEDAPAL has an extensive customer service center where registered clients can report their grievances via social media (Facebook, WhatsApp, Instagram, and Twitter), e-mail, phone, or video call. Residents, who are not yet registered as clients of SEDAPAL, cannot make use of this service, and therefore cannot report the issues they encounter while waiting to be registered and connected to the water infrastructure. Hence, although partly legible digitally, the informal city cannot materialize this limited legibility into the further development of the water infrastructure and remains illegible in the cadaster and the customer service center.

The geography of legibility, however, is not the mirror image of the formal city. As will be discussed in the following sections, each practice of legibility making adds a new layer of legibility, and therefore illegibility, to the image of the city. In practice, the digital legibility making of settlements adds to the distinction between formal neighborhoods (fully legible), informal neighborhoods that are digitally legible, and informal neighborhoods that are fully illegible. They are thereby creating a more differentiated view than the black and white model of the formal and the informal city.

4.2. Practices of legibility making of water distribution

The second practice of digital legibility-making is the centralized supervision of the water distribution system through the sectorization and the associated implementation of the supervisory control and data acquisition (SCADA) system in Lima's water infrastructure. In 1996, through a series of partially externally funded development projects, SEDAPAL started with the re-organization and division of the water distribution system into sectors to reduce the amount of non-revenue water and improve the efficiency of the distribution system. A sector is an area of around 2 km^2 containing between 2000 and 5000 households which can be connected or disconnected from the main pipelines independently. Today, the water distribution network is divided into more than 480 sectors. Each sector is connected to the main pipes with one valve through which the water pressure can be monitored and controlled. To fully operate, the sector needs to be hermetically sealed and equipped with functioning macro- and micrometers. The macro-meter measures the total amount of water that flows into the sector, and the micrometer measures the consumption of water per building or household. When fully controlled, the difference in the balance between the macro- and micrometers in the sector should be less than 25%. This target for the maximum ratio of non-revenue water (NRW) was set by the Japan International Cooperation Agency (JICA) and the World Bank as funding agencies in line with international benchmarks (JICA, 2011).

Figure 2. The SCADA control room at the La Atarjea water treatment plant in Lima. Source: (IDOM, 2015).

The sectorization of the water infrastructure aimed to improve water management in Lima by producing more accurate and timely data by sector. The measurements from the macro-meters are registered in a SCADA system and overseen from the central control room at the La Atarjea water treatment plant (see, Figure 2). As the overall supervision and control system, the SCADA system builds on and connects other geo-information systems in the operational branch of SEDAPAL. This allows SEDAPAL to control the water pressure from a distance, address interruptions, and schedule the water delivery. The division of the water distribution system into different sectors has become fundamental for the operational activities of SEDAPAL. It is not only the basis for their geo-information system detailing the flow of water in the city, but it also provides data about the quality of the pipes and allows for more efficient repair work. Driven by the wish to become more efficient, the system is divided into easily legible, and therefore, governable and administrable segments.

However, the legibility of the sectors is limited in two important ways. First, through the development and implementation of the sectoral system, SEDAPAL has attempted to anticipate future urban growth by already plotting sectors for areas, which had not yet been inhabited. However, urbanization in Lima does not always abide by the logic of the sectoral system. On the contrary, it responds to the implementation of technology in ways that were not anticipated. In practice, people often settle in physical proximity to already constructed water reservoirs, assuming this will increase their chances of being included in the future expansion of the piped infrastructure without being aware of the sectoral boundaries in place.

SEDAPAL makes a distinction between "controlled" sectors and sectors that are not yet operational. Lack of operationality can be due to technological failure or the incomplete installation of meters and valves to seal the area hermetically. Several sectors are still in the process of implementation, meaning that SEDAPAL has delineated the boundary within their geo-information system but not yet implemented the technology in practice. These sectors only exist in SEDAPAL's maps.

Second, there is a division between the supervisory system monitoring the sectors in the lower-lying areas, where water is circulated by gravity, and the city's areas at a higher altitude than La Atarjea (246 meters above sea level), in which pumps distribute the water. The two systems run in parallel, each monitoring a part of the water distribution system using different software and, most importantly, using a different frequency for data transmission. The SCADA system that collects data from the distribution system serviced by pumps depends on a public frequency that is often interrupted, losing the communication between the macro meters and the SCADA for months at a time. Contrastingly, the SCADA system for the part of the water distribution system driven by gravity has its own frequency and is, as a result, more reliable than the SCADA system for the pump-driven infrastructure. In general, the higher parts of the city, dependent on water from the pumps, are the newly constructed informal neighborhoods.

According to one of SEDAPAL's engineers, it was not a deliberate choice to use two SCADA systems in parallel. Rather, it resulted from the public procurement system in which the bid for the second system in the pump-supplied areas was won by a company using software that was not compatible with the already implemented SCADA for the gravity-based circulation. The effect is that the water distribution system in the low-lying neighborhoods is legible and, therefore, supervised in real-time. In contrast, the water distribution in the higher neighborhoods is often illegible and goes unsupervised for extended periods.

Legibility of the water distribution system is crucial as it informs SEDAPAL's decisions regarding the pressure and continuity of the water provided. First, since leakages are not detected quickly in the illegible sectors, water pressure is reduced to lessen the risk of high water loss. Secondly, to maintain governance over water distribution and its consumption in the illegible sectors, SEDAPAL rations the water for those areas. Only sectors that are fully legible receive water for 24 hours per day. We will return to this in the following section.

The sectorization and the SCADA system, as means for legibility making, help SEDAPAL monitor water distribution and possible leakages within the infrastructure, making it more efficient. However, people who live in sectors, where the technology is not fully implemented or where the SCADA system functions intermittently, are illegible or only sporadically legible. This differentiation in legibility of the water distribution system results in a higher and more regular water supply for those, who are legible, compared to people, who are only partially legible or illegible.

4.3. Practices of legibility making of water consumption

The third scale of legibility making in the water infrastructure is the micro metering of household water consumption. As of December 2019, 95.4% of Lima's piped water connections are metered (SEDAPAL, 2019). The purpose of metering water consumption per household is twofold. Primarily, the meters allow for more accurate billing of the water consumed. Secondly, the fine-grained data generated by the micro-meters gives a more comprehensive insight into patterns of water consumption within the city. Keeping track of consumption patterns can help with the prognosis of future water needs and identify possible issues in the water infrastructure such as leakages or clandestine consumption. Therefore, water meters have been an essential tool for SEDAPAL to reduce non-revenue water in their system.

Different types of micro-meters have been installed in Lima's households. Although the most recent meters can be read digitally, the majority of the meters are analog and read manually by an employee of SEDAPAL (INACAL, 2020). The data retrieved from the meters is then digitized and incorporated into the supervisory and control system of the administrative processes of SEDAPAL. Similar to the SCADA monitoring of the water distribution system's operation, this digital data technology monitoring the water consumption and payment is stooled on GIS.

Nevertheless, while the overall coverage of the consumption meters in Lima is relatively high, we see two critical limitations in making water consumption legible. First, measuring household water consumption with micro-meters is based on the assumption that residents live in individual households. In practice, specifically in Lima's older districts, many residents live in multi-family housing units, such as apartment buildings or "*quintas*", and rely on one entry point from which the water and the bill are divided among the tenants. SEDAPAL offers a guideline for the administration of the water amongst the tenants in which they suggest dividing the cost equally amongst each household, not taking into account the differences in consumption due to variations in the household size or water use. This requires a collective administration of the residents, and its success strongly depends on the community's organizational capacity. In these cases, the water consumption registered by SEDAPAL is that of the multi-family unit rather than the individual households. For the representation of people in the data, this is not insignificant as there can be up to 70 households within a *quinta*, and the number and height of apartment buildings in Lima is increasing.

Secondly, we see that the coverage of the meters within the city varies strongly from 41.3% in the least metered district to 99.9% in the most densely covered district (SEDAPAL, 2019). A partial explanation for this geographical difference lies in the lag in the implementation of meters by SEDAPAL in newly urbanized areas. Additionally, SEDAPAL argues that due to vandalism, the opposition of residents to the installation of the meters in their household, and the difficulty of planning the installation of meters in high-crime areas, some areas are under metered (SEDAPAL, 2014). Within public discourse, the meters are perceived as unreliable, even raising the suspicion that they function as a tool for SEDAPAL to raise water bills (Hoefsloot et al., 2020). In several central districts, neighborhoods have collectively opposed to installing the meters. One SEDAPAL employee explained they had ceased their efforts to implement the water meters in certain areas, since they had been harassed during their job. There are also cases reported, where communities removed the meters upon installation. As a result of this resistance, whole blocks remain illegible in the water consumption data.

This is important as SEDAPAL argues that only households with a water meter installed should continuously receive water. The argument for this is that households, which are metered, tend to use less water than those that are not, since they feel the financial consequences of water consumption directly in their monthly payments. The connections that do not have a meter are thus rationed and receive water only for a limited time per day, sometimes even as little as one hour.

Where meters have been installed, the digitalized data of the meter helps monitor whether the water consumption follows the expected pattern. Outliers in the data can signal various problems in the system. A high outlier can point toward an unreported

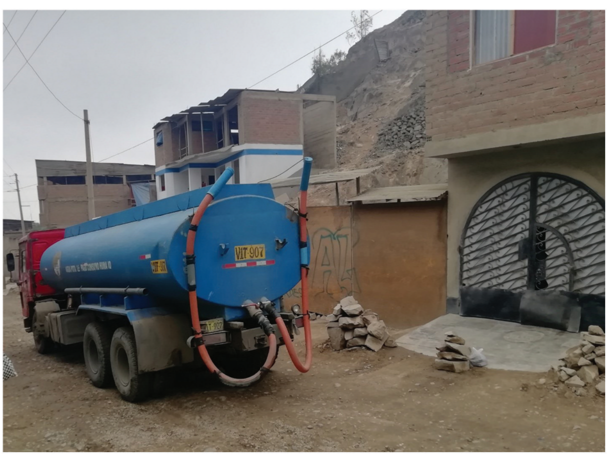

Figure 3. *Truck with water tank (camión cisterna) selling water to people and neighborhoods that are not connected to the piped water network. Photo was taken by first author. .*

leak or a sudden increase in consumption (as the result of a family visit, for example). Low outliers can signal that the meter is broken, the consumption has gone down drastically, or the metered pipe has been bypassed by constructing a clandestine connection. Engineers of SEDAPAL are using Google Street View and Google Maps to see if they can identify what the probable causes may be. For example, a construction site, an abandoned house, or buildings with rental advertisements can explain a reduction in metered water consumption. On the other hand, buildings with newly constructed stories or laundry drying on the rooftop, in combination with very little recorded consumption, might point toward a clandestine connection.

Thus, making water consumption legible is integral to SEDAPAL's strategy to reduce non-revenue water and to increase the infrastructure's efficiency. Nevertheless, in effect, it also provides SEDAPAL with arguments to ration the water delivery to households that are not metered. In addition, the data of the water meters in combination with Google Street View and Google Maps allows SEDAPAL to inspect changes in water consumption patterns on the household level, a much finer granularity than previously possible.

4.4. Water access while illegible

The nature of urban development in Lima illustrates how attempts of legibility making fail to deal with the speed and the uncertain and unpredictable nature of urban expansion and that the data created about the new settlements and water consumers of Lima excludes people and practices of accessing water resources. However, these illegible water consumers do deploy various strategies to access water. As described previously, they either buy water by tanker trucks (see, Figure 3), share connections with neighbors, or create clandestine connections to the pipelines. The water consumed through these alternative or clandestine methods is not or only poorly regulated and often sold at a higher price than the water delivered through the pipes. Importantly, as SEDAPAL is not the water vendor, the water quality is not monitored and hence not guaranteed by SEDAPAL.

While their water consumption is absent from the data used to manage the water infrastructure, it is present in the dominant narrative about water distribution in the city. In policy documents and master plans, their consumption is described as "illegal" or "clandestine" and grouped with other illegible water flows such as leakages and overflows under the category of non-revenue water (SEDAPAL, 2014; SUNASS, 2017).

As the discourse around the datafication of Lima's water infrastructure focuses on the reduction of non-revenue water, the people not counted are by association identified as illegal. On the ground, SEDAPAL is testing different strategies to detect illegible connections. In addition to using Google products to verify consumption patterns that are illogic according to the data, SEDAPAL is experimenting with the use of a ground-penetrating radar, or georadar, to survey the subsurface of the city for pipes. The georadar, installed on a small cart, can detect pipes and determine their type and width, making it possible to corroborate whether or not SEDAPAL has constructed the pipes. Any clandestine connections detected are removed, and the households connected will be fined.

Table 2. Summary of the identified types of water consumers that emerge from the datafication of the water infrastructure.

	Registered* and metered water consumer	Registered, non-metered water consumer	Alternative* water consumer	"Illegal"* water consumer	
Means of legibility making practices:	Of settlements: Cadaster and drone mapping	Yes	Yes	Pre-formalization	Pre-formalization
	In water distribution: Sectorization and SCADA systems	Yes	Yes/No	Yes/No	No
	Of water consumption: Micrometer, Google products, and Georadar	Yes	No	No	No/Yes**
Effects of (il)legibility on water distribution:	Water quality	Monitored	Monitored	Non-monitored	n/a
	Price	Pays a low rate for actual consumption	Pays a low rate for estimated consumption	Pays a high rate for actual consumption	n/a
	Reliability	High – continuous water supply	Low – rationed water supply	Low – rationed water supply	n/a

** Registered water consumers receive water from the public piped water infrastructure. Alternative water consumers often receive water from the tanker trucks or communal water taps. "Illegal" water consumers receive water via clandestine connections.*
***'Illegal' water consumers are strategically made legible by SEDAPAL using Google products and a georadar for the purpose of policing.*

5. Interpretation: Degrees of (il-)legibility and their effects on water provision and consumers

From the vantage point of SEDAPAL, looking through the data and following the legibility-making practices, we identified how the data image of the water infrastructure only partially represents the city. Legible are those parts of the city that are at one and the same time: mapped in SEDAPAL's cadaster, monitored in the SCADA system, and measured through the household water meters. However, on-the-ground reality, in which people receive water at irregular times, settle outside of pre-defined sector boundaries, and are not registered in the administrative databases, defies attempts to generate a fully comprehensive and standardized data image of the city. For every legibility-making practice that we discussed, we see that there is also illegibility created. These outcomes can be layered in different ways and, as such, create different degrees of legibility and illegibility in the hydrosocial geography across the city.

These findings align with other research that demonstrates how data are not fixed or objective but a link in a long chain of actions and decisions (Kitchin et al., 2016; Richter & Georgiadou, 2016). Specific about datafication is that (il)legibility making practices are layered; and people become categorized in multiple ways through a bundle of digital, technical, and spatial contingencies. In other words, the datafication – understood as the interplay between different legibility-making practices that are running partially concurrently – determines the relationship between SEDAPAL, as a state organization, and Lima's residents.

Table 2 summarizes how the interlinkages between the three scales of legibility-making practices analyzed in section 4 create a differentiated hydrosocial geography. That is, a geography characterized by distinct categories of water consumers as well as by varying, corresponding effects on the water delivery. Four types of water consumers in the city emerge from our analysis as a result of different degrees of (il)legibility: (i) the registered and metered water consumer, (ii) the registered, non-metered water consumer, (iii) the alternative water consumer, and (iv) the "illegal" water consumer (Table 2).

By unpacking the datafication of Lima's water infrastructure as layered legibility making practices, we see how the differentiation in water consumers is not the sole result of the unruliness of Lima's urban sprawl, the lack of control over water flows, or the unwillingness of residents to work within the imposed structure for water governance. Although all these elements play a part, the categorization of water consumers, as emerged from our analysis, is at the same time the result of the, sometimes arbitrary, choices made in the development of digital data technologies. A such, what appear to be purely technical interventions in the infrastructure and practices of legibility making are not independent of the broader political landscape (Amin & Thrift, 2017). The more comprehensive and detailed view of the water infrastructure makes it easier to control water flows within the city, one of the main goals of SEDAPAL. While at the same time, it has also provided the reasoning to deliver less water to illegible consumers and inspect unexpected water consumption patterns. Hence, as a result of the datafication, the water infrastructure has become more efficient to manage for SEDAPAL, more reliable for registered and metered water consumers, but less equitable for the residents at large. In other words, the direct impact of the datafication is the implicit and hierarchical categorization made between water consumers and the effects of this categorization are experienced through the differences in the quality, price and reliability of the water distributed. As illustrated in Table 2, the order created between water consumers in the data and urbanized and not-yet-urbanized areas exposes the interactions between (il) legibility, infrastructural development, and the relationships between the state and its residents (Sultana, 2020) and demonstrates how people living in conditions illegible by the state are subjected to illegalization as a governance strategy (Roy, 2018).

This in turn illustrates how legibility making, and illegibility making, are – sometimes strategic – parts of urban governance (Gandhi, 2017; Kalir & van Schendel, 2017). On the one hand, data can be created to make the inequality in water connections and consumption visible and accelerate the process of service provision. In addition, the exact enumeration of the people living in Lima in different degrees of informality and their water needs can serve as a tool to hold SEDAPAL accountable for the delivery of water services. On the other hand, as SEDAPAL has the mandate to connect all citizens of Lima to the water infrastructure, it can be strategic for SEDAPAL to maintain illegibility in areas that are difficult to connect to the infrastructure due to economic, political, or topographical reasons. At the same time, SEDAPAL uses Google products and the georadar selectively to make clandestine connections legible and police the people who consume water via those systems.

6. Conclusion

With the rapid implementation of digital data technologies in urban infrastructures, we are beginning to grapple with the effects of data on the distribution patterns of public services and resources. In this article, we aimed to answer how datafication, as layered legibility-making practices, shapes Lima's hydrosocial geography and affects residents as distinct water consumers. The case of Lima's water infrastructure interventions illustrates how implementing "smart" technologies and integrating multiple data sources for resource management increases the administrative order and control over urban resources. Yet, at the same time, datafication does not automatically lead to the further incorporation of residents into the water network, but produces new distinctions between spaces and between people. Moreover, this research argues how this act of differentiation structures the relationship between the water provider and consumer, and produces – more or less unintentionally – new categories of water consumers, which we have represented as four types.

Politically, the most significant distinction that emerges is that between the registered water consumers, on the one hand, and the deviant cases, the "illegal" water consumer, on the other. The crux lies in the normative notions attributed to the categories created (Roy, 2018). We see the implications of datafication in how it distinguishes between the consumers, who are legible, and those consumers, who are illegible and whose consumption is, in turn, criminalized and who become categorized as "illegal consumers" (Offenhuber, 2017; Pilo', 2017; Roy, 2018). In other words, datafication, especially through its – in this case at first sight implicit – categorizations – mediates people's claims and access to water resources, a key lesson that is similar to arguments made in research on Indian water infrastructure (Anand, 2017).

Thus, despite the rhetoric focusing on the benefits of data production for efficient infrastructural management, we observe that through a bundle of digital, technical, and spatial contingencies, (il)legibility-making practices create differential geographies in the city beyond the formal/informal dichotomy. Table 2 illustrates that the layering of legibility-making practices differentiates water consumers according to multiple characteristics and creates a gradient in legibility rather than a black and white image. This differentiation between these categories is not fixed, but continuously reconfigured through the production and bundling of data. And, importantly, through the strategic movement of people in and out of sight of administration and public service provision by the state.

The water infrastructure analyzed in this paper shows the value of approaching "smart city" developments as layered legibility-making practices. We found that in the "smart city," where various structured and unstructured sources of data come together and databases are made interoperable, it becomes increasingly important to consider not only the role of a variety of actors beyond the state that are making legible (Li, 2005; Taylor & Richter, 2017), but also the links between the variety of data sources and understand their interplay.

As newer streams of data and associated technologies interlink with existing practices and infrastructure, the longer-term effects on the distribution of urban resources remain dynamic and not entirely predictable. Future research will have to focus on how and where the top-down governance practices meet with daily everyday practices of people living in cities and consuming urban resources. This would reveal how the everyday experiences of people living in cities are processed and enacted through infrastructural development. As the case of Lima shows, this is specifically important in cities characterized by large socio-economic, spatial, and hydrological inequalities, where digital data technologies seek to increase equality and homogenization, but seem to introduce new differentiations, multiple layers of boundaries, and as such, reproduce the nature of inequality.

Acknowledgments

We want to express gratitude to the residents and community organizers in José Carlos Mariátegui, Barrios Altos, and Miraflores and the employees of SEDAPAL for their time, patience, and generosity in sharing their insights and knowledge for this research. Additionally, we thank the three anonymous reviewers for their valuable comments for improving this work. Finally, a special thanks goes out to Liliana Miranda Sara and Susana Gaete Sara from Foro Ciudades para la Vida for their generous help and guidance during the development of this research and in conducting fieldwork.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Fenna I. Hoefsloot http://orcid.org/0000-0002-3373-3580
Christine Richter http://orcid.org/0000-0002-2402-1768
Javier Martínez http://orcid.org/0000-0001-9634-3849
Karin Pfeffer http://orcid.org/0000-0002-6080-1323

References

Alda-Vidal, C., Kooy, M., & Rusca, M. (2018). Mapping operation and maintenance: An everyday urbanism analysis of inequalities within piped water supply in Lilongwe, Malawi. *Urban Geography*, *39*(1), 104–121. https://doi.org/10.1080/02723638.2017.1292664

Allen, A., Zilbert Soto, L., Wesely, J., Belkow, T., Ferro, V., Lambert, R., . . . Samanamú, A. (2017). From state agencies to ordinary citizens: Reframing risk-mitigation investments and their impact to disrupt urban risk traps in Lima, Peru. *Environment and Urbanization*, *29*(2), 477–502. https://doi.org/10.1177/0956247817706061

Amin, A., & Thrift, N. (2017). *Seeing like a city*. Polity Press.

Anand, N. 2017. *Hydraulic City*. Hydraulic City: Duke University Press. https://doi.org/10.1215/9780822373599.

Birnholtz, J. P., & Bietz, M. J. (2003). Data at work: Supporting sharing in science and engineering. In *Proceedings of the 2003 international ACM SIGGROUP conference on Supporting group work* (pp. 339–348). Sanibel Island, Florida, USA: ACM.

Bowker, G. C., & Star, S. L. (2000). *Sorting things out*. MIT Press.

Criqui, L. (2015). Infrastructure urbanism: Roadmaps for servicing unplanned urbanisation in emerging cities. *Habitat International*, *47*, 93–102. https://doi.org/10.1016/j.habitatint.2015.01.015

D'Ignazio, C., & Klein, L. F. (2020). *Data Feminism*. MIT Press.

El Comercio. (2019, September 26). *Marcha por el agua: Movilización contra privatización de Sedapal intenta llegar a San Isidro*. https://elcomercio.pe/lima/marcha-por-el-agua-movilizacion-contra-privatizacion-de-sedapal-intenta-llegar-a-ministerio-de-vivienda-en-san-isidro-noticia/

Elwood, S. (2008). Volunteered geographic information: Future research directions motivated by critical, participatory, and feminist GIS. *GeoJournal*, *72*(3–4), 173–183. https://doi.org/10.1007/s10708-008-9186-0

Fernández-Maldonado, A. M. (2008). Expanding networks for the urban poor: Water and telecommunications services in Lima, Peru. *Geoforum*, *39*(6), 1884–1896. https://doi.org/10.1016/J.GEOFORUM.2007.11.007

Gaffney, C., & Robertson, C. (2016). Smarter than smart: Rio de Janeiro's flawed emergence as a smart city. *City, Journal of Urban Technology*, *25*(3), 47–64. https://doi.org/10.1080/10630732.2015.1102423

Gandhi, A. (2017). The santioning state: Official permissiveness and prohibition in India. *Focaal - Journal of Global and Historical Anthropology*, *2017*(77), 8–21. https://doi.org/10.3167/fcl.2017.770120

Glasmeier, A., & Christopherson, S. (2015). Thinking about smart cities. *Cambridge Journal of Regions, Economy and Society*, *8*(1), 3–12. https://doi.org/10.1093/cjres/rsu034

Heeks, R., & Shekhar, S. (2019). Datafication, development and marginalised urban communities: An applied data justice framework. *Information Communication and Society*, *22*(7), 992–1011. https://doi.org/10.1080/1369118X.2019.1599039

Hoefsloot, F. I., Martínez, J., Richter, C., & Pfeffer, K. (2020). Expert-Amateurs and smart citizens: How digitalization reconfigures Lima ' s water infrastructure. *Urban Planning*, *5*(4), 312–323. https://doi.org/10.17645/up.v5i4.3453

IDOM. (2015). *Water control Peru*. Retrieved September 5, 2019, from https://www.idom.com/project/water-control-in-peru/#

INACAL. (2020). *Listado de Certificados de Aprobación de Modelo de Medidores de Agua*. Retrieved October 15, 2020, from https://www.inacal.gob.pe/repositorioaps/data/1/1/5/jer/aprobaciondemodelo/files/Listado_Modelos_Aprobados_Agua.pdf

INEI. (2019). *Perú: Formas de acceso al agua y saneamiento básico*. Instituto Nacional de Estadística e Informática.

Ioris, A. A. R. (2012). The neoliberalization of water in Lima, Peru. *Political Geography*, *31*(5), 266–278. https://doi.org/10.1016/J.POLGEO.2012.03.001

Ioris, A. A. R. (2016). Water scarcity and the exclusionary city: The struggle for water justice in Lima, Peru. *Water International*, *41*(1), 125–139. https://doi.org/10.1080/02508060.2016.1124515

JICA. (2011). *The preparatory survey on North Lima Metropolitan area water supply and sewerage optimization project (II) final report*. https://openjicareport.jica.go.jp/pdf/12047429_01.pdf
Kaika, M. 2012. *City of flows: Modernity, nature, and the city*. Routledge. https://doi.org/10.4324/9780203826928.
Kalir, B., & van Schendel, W. (2017). Introduction: Nonrecording states between legibility and looking away. *Focaal*, *2017*(77), 1–7. https://doi.org/10.3167/fcl.2017.770101
Karasti, H., Pipek, V., & Bowker, G. C. (2018). An afterword to "Infrastructuring and collaborative design. *Computer Supported Cooperative Work*, *27*(2), 267–289. https://doi.org/10.1007/s10606-017-9305-x
Kemerink-Seyoum, J. S., Chitata, T., Guzmán, C. D., Novoa-Sanchez, L. M., & Zwarteveen, M. Z. (2019). Attention to sociotechnical tinkering with irrigation infrastructure as a way to rethink water governance. *Water (Switzerland)*, *11*(8), 1–16. https://doi.org/10.3390/w11081670
Kitchin, R., & Dodge, M. (2017). The (In)Security of smart cities: Vulnerabilities, risks, mitigation, and prevention. *Journal of Urban Technology*, *26*(2) , 1–19. https://doi.org/10.1080/10630732.2017.1408002
Kitchin, R., Lauriault, T. P., & McArdle, G. (2015). Knowing and governing cities through urban indicators, city benchmarking and real-time dashboards. *Regional Studies, Regional Science*, *2* (1), 6–28. https://doi.org/10.1080/21681376.2014.983149
Kitchin, R., & Lauriault, T. P. (2018). Toward critical data studies : Charting and unpacking data assemblages and their work. In J. Thatcher, J. Eckert, A. Shears (Eds.), *Thinking big data in geography: New regimes, new research*. University of Nebraska Press. https://doi.org/10.2307/j.ctt21h4z6m
Kitchin, R., Maalsen, S., & McArdle, G. (2016). The praxis and politics of building urban dashboards. *Geoforum*, *77*, 93–101. https://doi.org/10.1016/J.GEOFORUM.2016.10.006
Lambert, R., & Allen, A. (2016). Participatory mapping to disrupt unjust urban trajectories in Lima. In P. Imperatore (Ed.), *Geospatial technology - Environmental and social applications* (pp. 143–165). InTech.https://doi.org/10.5772/64303
Leonard, J. B. (2000). City profile: Lima. *Cities*, *17*(6), 433–445. https://doi.org/10.1016/S0264-2751(00)00047-0
Li, T. M. (2005). Beyond "the State" and Failed Schemes. *American Anthropologist*, *107*(3), 383–394. https://doi.org/10.1525/aa.2005.107.3.383
Lindley, S. E., Thieme, A., Taylor, A. S., Vlachokyriakos, V., Regan, T., & Sweeney, D. (2017). Surfacing Small Worlds through Data-In-Place. *Computer Supported Cooperative Work: CSCW: An International Journal*, *26*(1–2), 135–163. https://doi.org/10.1007/s10606-017-9263-3
Luque-Ayala, A., & Marvin, S. (2015). Developing a critical understanding of smart urbanism? *Critical Commentary Urban Studies*, *52*(12), 2105–2116. https://doi.org/10.1177/0042098015577319
Luque-Ayala, A., & Marvin, S. (2016). The maintenance of urban circulation: An operational logic of infrastructural control. *Environment and Planning D: Society and Space*, *34*(2), 191–208. https://doi.org/10.1177/0263775815611422
Luque-Ayala, A., & Neves Maia, F. (2019). Digital territories: Google maps as a political technique in the re-making of urban informality. *Environment and Planning D: Society and Space*, *37*(3), 449–467. https://doi.org/10.1177/0263775818766069
Mayer-Schönberger, V., & Cuckier, K. (2013). *Big Data: A revolution that will transform how we live, work, and think*. John Murray Publishers.
McFarlane, C., & Söderström, O. (2017). On alternative smart cities. *City*, *21* (3–4), 312–328. http://www.tandfonline.com/action/journalInformation?journalCode=ccit20.
Miranda Sara, L., Jameson, S., Pfeffer, K., & Baud, I. (2016). Risk perception: The social construction of spatial knowledge around climate change-related scenarios in Lima. *Habitat International*, *54*(2), 136–149. https://doi.org/10.1016/J.HABITATINT.2015.12.025
Offenhuber, D. (2017). *Waste Is Information: Infrastructure Legibility and Governance*. The MIT Press.
Pelizza, A. (2017). Disciplining change, displacing frictions two structural dimensions of digital circulation across land registry database integration. *TECNOSCIENZA: Italian Journal of Science & Technology Studies*, *7* (2), 35–60. http://www.tecnoscienza.net/index.php/tsj/article/view/272/176

Pfeffer, K. (2018). *Knowing the city*. University of Twente.
Pilo', F. (2017). A socio-technical perspective to the right to the city: regularizing electricity access in Rio de Janeiro's Favelas. *International Journal of Urban and Regional Research*, *41*(3), 396–413. https://doi.org/10.1111/1468-2427.12489
Rabari, C., & Storper, M. (2015). The digital skin of cities: Urban theory and research in the age of the sensored and metered city, ubiquitous computing and big data. *Cambridge Journal of Regions, Economy and Society*, *8*(1), 27–42. https://doi.org/10.1093/cjres/rsu021
Republica del Perú. Decreto supremo N° 029-2019-VIVIENDA. (2019). *Ministerio de Vivenda, Construcción y Saneamiento*. https://www.gob.pe/institucion/vivienda/normas-legales/354318-029-2019-vivienda
Richter, C., & Georgiadou, Y. (2016). Practices of legibility making in Indian cities: Property mapping through geographic information systems and slum listing in government schemes. *Information Technology for Development*, *22*(1), 75–93. https://doi.org/10.1080/02681102.2014.886548
Riofrío, G. (2003). Pobreza y desarrollo urbano en el Perú. In C. Eduardo Aramburú, E. Ballón, J. Barreda, M. Llona, A. Huerta Mercado, D. Ramírez Corzo, G. Riofrío, P. Vega Centeno, M. Zolezzi (Eds.),) *Las ciudades en el Perú*, Perú Hoy (pp. 71–111). DESCO.
Roy, A. (2018). The potency of the state: Logics of informality and subalternity. *Journal of Development Studies*, *54*(12), 2243–2246. https://doi.org/10.1080/00220388.2018.1460470
Scott, J. C. (1999). *Seeing like a state - How certain schemes to improve the human condition have failed*. Yale University Press.
SEDAPAL. (2014). *Plan Maestro de Los Sistemas de Agua Potable y Alcantarillado. Tomo I - Diagnostico Volumen I*. Gerencia de Desarrollo e Investigación.
SEDAPAL. (2015). *Plan Estratégico de Tecnologías de la Información y Telecomunicaciones de Sedepal*. Resolución de Gerencia General. Servicio de Agua Potable y Alcantarillado de Lima.
SEDAPAL. (2018). SCADA. Retrieved November 3, 2018, from http://www.sedapal.com.pe/scada
SEDAPAL. (2019). *Conexiones con medidor y sin medidor - catastro de conexiones por estados del suministro*. Equipo Gestion Comercial y Micromedicion.
Shaw, J., & Graham, M. (2017). An informational right to the city?. In J. Shaw, M. Graham (Eds.), *Our digital rights to the city* (pp. 4–5). Meatspace Press.
Sultana, F. (2020). Embodied intersectionalities of urban citizenship: Water, infrastructure, and gender in the global south. *Annals of the American Association of Geographers*, *110*(5), 1407–1424. https://doi.org/10.1080/24694452.2020.1715193
SUNASS. (2017). *Sunass advierte sobre riesgo de cárcel a personas con conexiones clandestinas*. https://www.sunass.gob.pe/doc/NotasPrensa/2017/Septiembre/np271_2017.pdf
Swyngedouw, E. N. Heynen, M. Kaika*et al* and E. Swyngedouw. (2006). In N. Heynen, M. Kaika, and E. Swyngedouw (Eds.), *The nature of cities: Urban political ecology and the politics of urban metabolism* (p. 288). Routledge. https://doi.org/10.2747/0272-3638.28.2.206
Taylor, L., & Broeders, D. (2015). In the name of development: Power, profit and the datafication of the global South. *Geoforum*, *64*, 229–237. https://doi.org/10.1016/j.geoforum.2015.07.002
Taylor, L., & Richter, C. (2017). The power of smart solutions: Knowledge, citizenship, and the datafication Of Bangalore's water supply. *Television and New Media*, *18*(8), 721–733. https://doi.org/10.1177/1527476417690028
Truelove, Y. (2018). Negotiating states of water: Producing illegibility, bureaucratic arbitrariness, and distributive injustices in Delhi. *Environment and Planning D: Society and Space*, *36*(5), 949–967. https://doi.org/10.1177/0263775818759967
van Dijck, J. (2014). Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology. *Surveillance and Society*, *12*(2), 197–208. https://doi.org/10.24908/ss.v12i2.4776
Vanolo, A. (2014). Smartmentality: The smart city as disciplinary strategy. *Urban Studies*, *51*(5), 883–898. https://doi.org/10.1177/0042098013494427
Verrest, H., & Pfeffer, K. (2018). Elaborating the urbanism in smart urbanism: Distilling relevant dimensions for a comprehensive analysis of Smart City approaches. *Information Communication and Society*, 1–15. https://doi.org/10.1080/1369118X.2018.1424921

Wesselink, A., Kooy, M., & Warner, J. (2017). Socio-hydrology and hydrosocial analysis: Toward dialogues across disciplines. *Wiley Interdisciplinary Reviews: Water*, *4*(2), e1196. https://doi.org/10.1002/wat2.1196
World Bank. (2011). *Optimization of Lima water and sewage systems*. http://documents.worldbank.org/curated/en/906631468059355803/pdf/Integrated0Saf1et010Appraisal0Stage.pdf
World Bank. (2018). *Modernization of water supply and sanitation services project*. http://documents1.worldbank.org/curated/en/118971532835034687/pdf/Peru-Modernization-PAD-07092018.pdf
Wutich, A., Beresford, M., & Carvajal, C. (2016). Can informal water vendors deliver on the promise of a human right to water? Results from Cochabamba, Bolivia. *World Development*, *79*, 14–24. https://doi.org/10.1016/j.worlddev.2015.10.043

Appendix 1: Codebook used in qualitative analysis

This codebook contains the list of used for the qualitative analysis of the data, texts, notes, and transcripts in ATLAS.ti™. Please note that some codes have been made prior to the analysis and other were created during the coding process as new themes emerged. As noted in section 3 of this research, the coding was conducted by the first author, who also collected the data in the field.

Code	Description
Actors involved	This code should give an overview of the different actors involved in the development of the project.
Data strategy	This code relates to text that described how data in itself is strategic or how data is used strategically.
Data production	This code is used for text which describes how data about the water infrastructure is produced in Lima.
Data use	This code relates to text that describes how data is used in urban water management. Data sharing amongst institutions is also a part of data use and thus coded in this code.
Description of SCADA	The text related to this code should explain the design, implementation, and/or functioning of the SCADA system.
Start/End date	This the date (year) that the infrastructural project has started and (is supposed to be) finished.
Functioning of system	The text related to this code should explain the design, implementation and/or function of the overall water infrastructure. Specifically, this refers to the water distribution system.
Funding agency	This is the main financer of the infrastructural project.
Hardware of information system	This code relates to the physical materials used in the (development) of the digital information system.
Institutional relationships	This code relates to text describing how different institutions involved in water management are related to each other/work together.
Justification for solution proposed in project	The text related to this code should answer the question: How is the proposed solution to the stated problem justified in the text? Why do they choose this solution as opposed to others?
Proposed solution to stated problem	The text related to this code should answer: what solution is proposed to solve the problem as stated in the text? For example, the stated problem is: a neighborhood is not connected to the water pipes. The proposed solution is: We need to expand the water pipes to include that neighborhood.
Sectorization	This code is related to the text that describes the process of sectorization of the water infrastructure in Lima.
Stated problem that is to be solved	The text related to this code should answer: what is the problem that is identified in the text that needs to be solved? For example, the stated problem is: a neighborhood is not connected to the water pipes.
Targeted location of project	Which neighborhoods of the city does the project focus on/take place?
Urban development process	This code relates to text that describes the urban expansion and growth process in Lima.
Sedapal – Administration	Describes the administration of the water infrastructure by SEDAPAL
Sedapal – Construction	Describes the construction of the water infrastructure by SEDAPAL