ISSN 2564-3843
DOI 10.26443/glsars.v1i1.120

MAKING DATA VISIBLE IN PUBLIC SPACE*

Abstract

"Transparency" is continually set as a core value for cities as they digitalize. Global initiatives and regulations claim that transparency will be key to making smart cities ethical. Unfortunately, how exactly to achieve a transparent city is quite opaque. Current regulations often only mandate that information be made accessible in the case of personal data collection. While such standards might encourage anonymization techniques, they do not enforce that publicly collected data be made publicly visible or an issue of public concern. This paper covers concerns for data transparency in public space. The first, why data visibility is important, sets the stage for why transparency cannot solely be based on personal as opposed to anonymous data collection as well as what counts as making data transparent. The second concern, lessons and shortcomings of current regulations and initiatives, addresses present challenges creating public space that communicates its sensing capabilities without overwhelming the public. The final section, what regulations are necessary for data visibility, argues that for public space to be transparent the government needs to step in to regulate standardized signaling, a sensor registry, contextual data accessibility and increased data literacy education.

Key Words

Smart city, transparency, visibility, signage, data Awareness

* **Sage Cammers-Goodwin**; PhD Candidate, University of Twente and **Naomi van Stralen**; MSc Candidate, University of Twente. Contact: s.i.cammers-goodwin@utwente.nl and naomivanstralen@gmail.com.

Résumé

La « transparence » est présentée comme une valeur essentielle pour les villes entreprenant un processus de numérisation. Les initiatives et réglementations mondiales affirment que la transparence est la clé pour résoudre les défis éthiques présentés par les villes intelligentes. Malheureusement, le chemin exact pourcréer une ville intelligente et transparente demeure nébuleux. En effet, les réglementations actuelles se limitent souvent à l'accès à l'information lors de collectes de données personnelles. Bien que ces normes puissent encourager les techniques d'anonymisation, elles n'adressent pas les questions de visibilité et autres préoccupations publiques.

Cet article traite des préoccupations relatives à la transparence des données dans l'espace public. La première de ces préoccupations – l'importance de la visibilité des données – permet d'expliquer pourquoi la transparence ne peut pas reposer uniquement sur la collecte de données personnelles, par opposition à la collecte de données anonymes. Elle pose également les bases d'une définition de la transparence des données. La deuxième préoccupation – les leçons et les lacunes des réglementations et initiatives actuelles – permet d'analyser les défis actuels quant à la création d'un espace public usant de ses capacités de détection sans pour autant accabler le public. La dernière partie de cet article traite de la visibilité et soutient que pour que l'espace public soit transparent, le gouvernement doit intervenir pour réglementer une signalisation normalisée, un registre de capteurs, l'accessibilité des données contextuelles et une meilleure éducation.

Mots-clés : Ville intelligente, transparence, visibilité, signalisation, sensibilisation aux données (*data awareness*)

BACKGROUND

Cities have quietly grown smarter. While some cities proudly brand themselves as smart cities, data collection in public space is still shrouded in mystery for most citizens. Smart street lights can disguise themselves as "normal" streetlights. Artificial Intelligence (AI) assisted security cameras may be mistaken for basic cameras. Moreover, information on where to access the data collected by these internet of things (IoT) technologies,[1] who is collecting the information, and for what purpose may be even more hidden than the sensor itself. There is a growing effort in some cities to confront the lack of transparency over data collection in public space. Such efforts are commendable, but there is still work to be done to create regulation that ethically ensures transparency, especially when the metrics that determine transparency may ironically be unclear. This paper tackles the conundrum of data visibility in public space by outlining *why* transparency is important, *how* data collection transparency works in practice, and lastly *what* regulations are necessary to ensure that data is effectively transparent.

Transparency has become a sought-after goal as cities continue to digitalize. Multiple international smart city initiatives name "transparency" as a main tenet of responsible development. TADA, a Dutch initiative for responsible smart cities, lists transparency as number five out of six necessities on their manifesto.[2] Similarly, Cities for Digital Rights, an initiative

[1] The Internet of things describes sensor embedded technologies that can communicate with other devices whether it be through the internet or Bluetooth connections.

[2] 'The 6 principles of our manifesto' (*TADA*) <https://tada.city/en/home-en/> accessed 4 May 2021

created in partnership with Amsterdam, Barcelona and New York in late 2018, includes principles such as transparency, universal internet access, and data literacy.[3] Effort has been made to standardize IoT signage in public space by Digital Trust for Places and Routines (DTPR), a now independent initiative started by Alphabet Inc's Sidewalk Labs Toronto Waterfront project.[4] Legally, transparency is an underlying thread in data regulations worldwide. The right to access information about personal data collection is present in laws in the European Union, Brazil, Canada, China, Thailand, Australia, Japan, South Korea, Chile, New Zealand, India, South Africa, and the State of California.[5] Amsterdam plans to activate a mandatory public data register in October 2021.[6]

While not the first internet-based privacy legislation, the European Union's 2018 General Data Protection Regulation (GDPR) arguably has had the strongest impact on global privacy legislation due to the combined power of its member countries.[7] Succeeding privacy laws, such as

---

[3] Since conception membership has grown. They now have over 46 member cities worldwide. 'Declaration of Cities Coalition for Digital Rights' (*Cities for Digital Rights*) <https://citiesfordigitalrights.org/declaration> accessed 4 May 2021; 'Here you will find cities that already joined the initiative to protect, promote and monitor residents' and visitors' digital rights.' (*Cities for Digital Rights*) <https://citiesfordigitalrights.org/cities> accessed 6 May 2021.

[4] 'DTPR is an open-source communication standard that enables transparency, accountability, and control for people. The standard is free for anyone to use.' (*DTPR)* <https://dtpr.helpfulplaces.com/> accessed 21 April 2021. DTPR formerly stood for 'Digital Transparency in the Public Realm' but has since changed its name since leaving Alphabet Inc. <https://github.com/helpful-places/dtpr> accessed 11 November 2021

[5] Dan Simmons '12 Countries with GDPR-like Data Privacy Laws' (*Comforte*, 12 January 2021) <https://insights.comforte.com/12-countries-with-gdpr-like-data-privacy-laws> accessed 4 May 2021.

[6] 'Inspraak Verordening meldingsplicht sensoren' (*Gemeente Amsterdam,* 24 February 2021*).* <https://bekendmakingen.amsterdam.nl/bekendmakingen/publicatie/inspraak/inspraak-sensoren/> accessed 7 May 2021.

[7] Korea and Japan's privacy regulations were enacted in 2011 and 2017 respectively before the 2018 European Union General Data Protection Act (GDPR). Simmons (n 5).

Brazil's *Lei Geral de Proteção de Dados* (LGPD) were openly modeled after the EU regulation. One of GDPR's main goals is to empower individuals to maintain their "fundamental rights and freedoms" over their personal data.[8] Although this regulation limits the ability to collect data without informing subjects in advance, it fails to clarify *how* information should be presented in the built environment. Furthermore, the regulation focuses solely on personal data. While regulations such as GDPR set a precedent for increasing awareness and consent regarding personal data collection, they fall short of ensuring a fully transparent smart public space.

Local efforts have also been made to specifically address algorithm transparency in public space such as Amsterdam and Helsinki's 2020 initiative to register all public AI decision making.[9] Both of these registers are operated by Saidot, a for profit company with the focus of creating more transparent AI systems. These registries are still under development, but one can already visit and learn more about algorithmic decision making in both cities.[10] Tools like these provide increased accountability for government algorithms, but may leave out corporate and academic projects. There is also a disconnect from the algorithm to the visibility of the data collection in the built environment. One can interact with IoT infrastructure and not be aware of the AI registry or the fact the infrastructure is "smart."

---

[8] General Data Protection Regulation [2016] *OJ L 119*.

[9] Meeri Haataja, Linda van de Fliert and Pasi Rautio, 'Public AI Registers' (*White Paper*, 2020) <https://uploads-ssl.webflow.com/5c8abedb10ed656ecfb65fd9/5f6f334b49d5444079726a79_AI%20Registers%20-%20White%20paper%201.0.pdf> accessed 7 May 2021.

[10] City of Amsterdam Algorithm Register Beta (*Gemeente Amsterdam*) <https://algoritmeregister.amsterdam.nl/> accessed 4 May 2021; City of Helsinki AI Register (*Helsinki*) <https://ai.hel.fi/en/ai-register/> accessed 4 May 2021.

One of the roadblocks to mandating a transparent public space is that there is no specific way to achieve it. Synonyms to "transparency" include "frank," "obvious," and "understandable." Applying these notions to the built environment means that IoT infrastructures will need to be able to speak for themselves in a way that the public can understand. Transparency requires communication between a listener and a speaker. Detailed signs could be placed on every sensor in public space, but this would be useless if people could not read or understand the language used. Alternatively, the public could be educated to recognize cues from the infrastructures in public space to know what sensors are installed and where to go to access more information. Both of these avenues would make data collection transparent.

Klaus Schwab, founder and executive chairman of the World Economic Forum, argues that depending on the policies of the government, cities can fully capitalize on the current technological revolution or be left behind in development.[11] In the rush to take advantage of new digital solutions, it is important to ensure that the promise of new technology does not undermine the autonomy of the public. The following sections combine knowledge from the fields of ethics of technology and design research to build benchmarks for data visibility in public space. Additionally, this paper uses a smart bridge, printed by metal 3d-printing company MX3D, to pragmatically address the issue of data transparency in public space. (See Figure 1). As part of this research, an ethics committee approved survey was conducted to find out what individuals found most important to know about the bridge data collection onsite.[12]

---

[11] Klaus Schwab, *The Fourth Industrial Revolution* (1st edn, World Economic Forum 2016) 73 <https://law.unimelb.edu.au/__data/assets/pdf_file/0005/3385454/Schwab-The_Fourth_Industrial_Revolution_Klaus_S.pdf>.

[12] The survey reached 32 participants. Out of these participants 59% of the participants had seen the bridge before, 72% of the people would like to know sensors are present on the bridge if they would have to cross

Figure 1. Sensors under the MX3D smart bridge (source: Authors)

Since completion in 2018, the MX3D bridge has been affixed with sensors for primarily digital twin studies.[13] Partners involved in the sensor work are multinational and include both academic and commercial institutions. The bridge was placed in Amsterdam's Red Light District summer 2021. Due to slow bureaucratic processes to approve the sensors, structural integrity of the bridge, physical location permits, and canal wall durability, placement was delayed for over two years. In order to decrease hacking risks, it was decided that the sensor data would be ported via ethernet cables to a local server and from there be securely transferred to long term cloud storage. Even though the bridge is installed in a self-proclaimed smart city focused municipality,[14] in practice,

---

it, 88% would want to know the purpose of the data collection on location, 53% would want to know why personal data in specific is collected on location, 63% would want to know where to find more information.

[13] It is fitting that the bridge is built to have a digital twin, given that its placement city of Amsterdam hopes to be a digital twin city. See: 'Digital Twin voor stedenbouw' (6 December 2019) <https://storymaps.arcgis.com/stories/71cd7ef2f092419ca1f02415e35d7d2f> accessed 4 May 2021.

[14] Renata Paola Dameri, *Smart city: How to create public and economic value with high technology in*

navigating the data visibility system has been challenging for the project partners (including the authors).

Nevertheless, the past two and a half years working with the bridge has given us collateral knowledge as well as research by design insights relating to three main concerns on increasing data visibility and transparency in public space. The first concern, 'why data visibility is important,' helps build metrics for good legislation. The second concern, 'the challenges of data visibility in practice,' illuminates current roadblocks and solutions for data transparency. And finally, 'regulations for data visibility', accounts for bureaucratic shortcomings to create a functional system.

PART 1. WHY DATA VISIBILITY IS IMPORTANT

Arguably, the current primary function of signage for personal data collection in public space is to prevent bad actors from performing criminal behavior or to meet the bare minimum for privacy legislation. Fully informing the public as to what sensors are in public space, who is controlling them, and why they are there – information that may increase individual autonomy – is not a top concern.[15] Signage is more often used as a tool to limit agency as opposed to increasing it. For example, if a city wants drivers to slow down it may invest in speed monitors that the driver

urban space (Camille Rosenthal-Sabroux eds, CH Comparing smart and digital city: Initiatives and Strategies in Amsterdam and Genoa. Are they digital and/or smart?, Springer International Publishing 2014) 45.

[15] In "The Internet of Us" Michael Lynch connects the concept of privacy to autonomy. He argues that the main concern of lack of transparency in data collection is that it undermines individual autonomy. This ends up becoming a power concern when individuals are used as a means to an end by corporations and governments. Michael Patrick Lynch, *The Internet of Us* (W.W. Norton, Liveright Publishing Corporation 2016).

themselves can read. If a city wants to ticket bad actors it may introduce speed cameras without any indication of an IoT device except for the camera itself. Signage is then primarily used as a nudging tool as opposed to an indicator of what sensors are affixed in public space.[16] In reality any form of informing the public cannot be completely transparent because truth is translated through a sign.

In the case of the smart bridge, it would be easy to ignore data visibility. Given that privacy regulations such as the General Data Protection Regulation (GDPR) only apply to personal data, anonymous data is freely open for collection and processing without notification.[17] Personal data, or data that can be traced back to the individual, may warrant a small sign, such as a diagram of a camera, which at best might inform those on the lookout for such signage. On the bridge, the long-term cloud server should not store anything that could be classified as personal. Figure 2 shows what data the bridge will be collecting. Figure 3 shows that only anonymized data will be saved in long-term storage from camera footage. (Software will run over raw camera footage to "skeletonize" individuals walking over the bridge.) Figure 4 shows data collected from one of the bridge's accelerometers.

[16] Nudging is not necessarily a bad thing as it can help people do things that are good for them as a sort of soft paternalism. Nonetheless, a nudge might have different goals than an autonomy boosting sign on how the IoT system is used. See for more information on nudging: Cass R. Sunstein. "Nudging: A Very Short Guide" [2014] 37 J. Consumer Pol'y 583.

[17] Point 26 of the introductory section of the General Data Protection Regulation states that anonymized data is not subject to the principles of data protection. The data protection is active on identifiable data, including pseudonymized data.

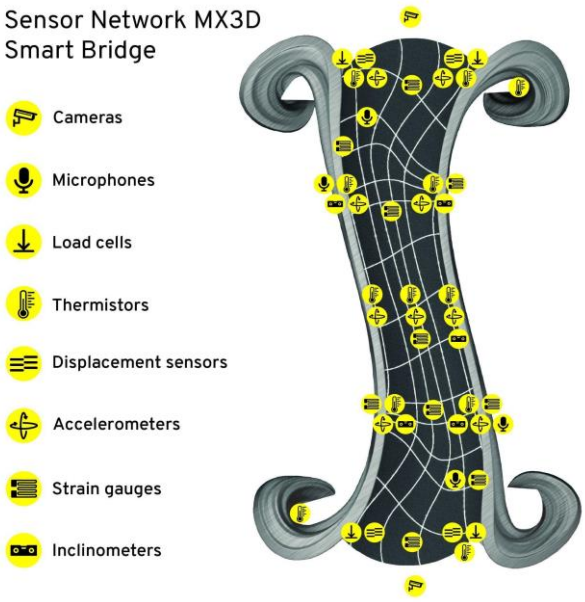

Figure 2. A simplified overview of the sensor network on and around the MX3D bridge. Microphones were discontinued (source: Authors)

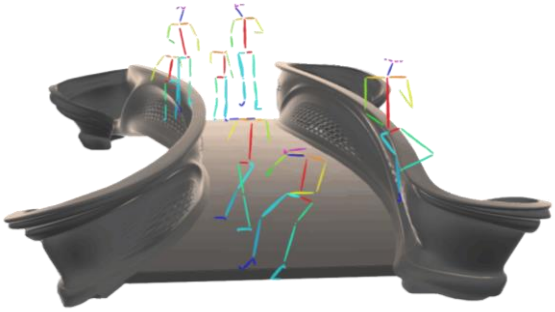

Figure 3. Skeletonized camera footage[18]

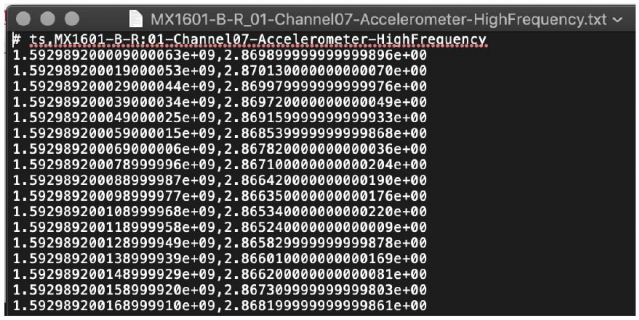

Figure 4. Accelerometer data from the bridge (source: Authors)

[18] Image skeletonized data, edited. Kean Walmsley, 'Realtime visualization of sensor data from the MX3D bridge' (24 October 2018) <https://www.keanw.com/2018/10/realtime-visualization-of-sensor-data-from-the-mx3d-bridge.html> accessed 5 May 2021.

Nonetheless, it is possible that data be traced back to an individual via changes in the data storage, increased machine learning capabilities, or combining data sources. For example, it was planned for the bridge to store microphone data every 30 seconds across a few frequencies to measure loudness in the area. While this data is anonymous, one could also use microphones to record continuous sound. Alternatively, with increased machine learning capabilities it might become easier to identify people through walking IDs based on their weight and movement pattern. External data such as metro cards and LinkedIn pages theoretically could be connected with the bridge data to predict individuals who crossed the bridge at a certain time to get to work.

In order to ethically bring the internet of things (IoT) into cities, it is important to understand the reasons why data visibility is important. Without such a foundation it is unlikely that legislation to inform the public will be promulgated and enacted. Further, tools to increase data visibility such as those discussed in the next section are unlikely to be employed if the public remains unaware. Government often operates under a utilitarian mindset to bring the greatest good (often economically) to the preferred members of its closed society.[19] Even though it may be tempting when introducing IoT devices in public space to primarily focus on utilitarian consequentialist justifications – such as advancing structural health monitoring and developing new machine learning techniques through a smart bridge – it is equally important to consider individual autonomy of those interacting with IoT infrastructure. If members of the public are considered moral agents, who are eligible to vote, run for office, pay taxes, and participate in

[19] Riley, Jonathan. 'Utilitarian Ethics and Democratic Government.' (1990) 100(2) Ethics 335. JSTOR <www.jstor.org/stable/2381000> accessed 15 Apr. 2021.

society, then they should also be made aware that their behavior, anonymous or otherwise, may be converted to data, stored, analyzed, and used by various parties.

Safety, security, and sustainability of the community are common justifications for IoT devices.[20] (Ironically these same values can be used as arguments against IoT devices as they create new hackable dependencies.)[21] While the safety of the group may outweigh the personal freedom of an individual, it is unclear what level of effectiveness outweighs other concerns such as individual freedom, agency, and self-governance. Consequentialist ethics allows one to take a deontologically questionable decision and justify it due to possible potential outcomes. Speed cameras that take pictures of vehicle license plates become justified because only wrongdoers may get punished and their punishment theoretically maximizes wellbeing for the greater society. People who do not know the camera is there and equally who do not break the law, will not get hurt. This creates a loophole to bypass the consideration of informing the public entirely. A similar concern might occur for IoT devices designed to measure "normal" behavior, such as WIFI-counting devices that measure movement inside a region though smartphone connectivity. The argument is that the technology is not going to hurt anyone. Moreover, if people were informed they would act differently or become worried, so the negative consequences of informing outweigh the positives of letting people know. This has been a subject of debate in the smart bridge research group. Is it possible that transparency may cause undue harm?

---

[20] Michelle Cayford & Wolter Pieters 'The effectiveness of surveillance technology: What intelligence officials are saying' (2018) 34(2) The Information Society 88, DOI: 10.1080/01972243.2017.1414721.

[21] Kevin Fu and others, 'Safety, Security, and Privacy Threats Posed by Accelerating Trends in the Internet of Things.' (*White Paper*, 2017) <http://cra.org/ccc/resources/ccc-led-whitepapers/> accessed 7 May 2021.

In truth, the sensored environment is becoming the new normal. It is a fallacy to believe that IoT left undercover will capture the natural citizen.[22] The best an IoT device can do is capture the "natural" uninformed citizen along with IoT literate citizens, ranging from the possibly guarded to those who may not care. Those left unaware of IoT advancements are stripped of the freedom to react to the devices. It is "natural" for humans to interact with objects in public space. Disguising a sensor may provide the most benefit to the informed, whether that be the government, corporations, or potential hackers. This is not to say that it is ethical to overwhelm users of public space by exposing them to a stressful environment. If the sensors in public space actually do provide a benefit to the public then awareness of them should not cause concern. However, if sensors do pose a threat to members of society, disguising IoT devices to subdue possible public response is not only dishonest, it limits individual agency and autonomy.

There are additional shortcomings to the "not going to hurt anyone" perspective. One being that people are notoriously bad at predicting the future. In the case of the embedded smart bridge it is unclear what machine learning capabilities will be after the bridge data has long been collected. With an increasing volume of datasets collected both in public and private space, it is also uncertain how many datasets could become synched and transform anonymous data into personal data. In April 2021, the city of Enschede, the Netherlands was retroactively fined €600,000 for the period they had their WIFI-tracking service active because it was found that on

[22] The state of being surveilled has its own phenomenological experience (see citation). Moreover, if one acts as if they are not being surveilled in a data gathering environment then they are not acting in accordance to how they would act with full information. In this light the behavior could be deemed fake. Kirstie Ball (2009) EXPOSURE, Information, Communication & Society, 12:5, 639-657, DOI: 10.1080/13691180802270386 https://www.tandfonline.com/doi/full/10.1080/13691180802270386.

low traffic days individuals could be tracked. The city is currently fighting the lawsuit on grounds that they never used the services in this theoretically privacy infringing manner.[23]

Furthermore, society is structured unequally. Those in government often do not overlap with the underprivileged. There are individuals whose identities whether it be disability, gender, sexuality, race, or age make them flags for state sanctioned discrimination.[24] These are the people who are often accidentally harmed because they are not the majority group considered in the consequentialist mindset. Consider for example if data from the bridge is used later for decision making processes and wheelchair users are not accurately labeled in the dataset. Minority populations are also often last to be informed about IoT devices in public space because they are not the target audience of government or corporations. Not informing everyone, even about anonymous data collection, actually might have unforeseen consequences for some.[25]

To be clear, the argument for data visibility is not a debate over whether or not IoT should be embedded in public space – that is a different question entirely, which this paper makes no attempt to address. It is also not a question of data ownership, although lessons from the visibility debate do run parallel to that discussion given that providing access to data is one form of data transparency. The argument for data visibility simply follows that people should be reasonably

---

[23] Privacywaakhond legt Enschede boete op van 600.000 euro vanwege wifitracking (*NOS*, 29 April 2021) <https://nos.nl/artikel/2378665-privacywaakhond-legt-enschede-boete-op-van-600-000-euro-vanwege-wifitracking> accessed 5 May 2021.

[24] Cammers-Goodwin, Sage. (Forthcoming 2021). Revisiting Smartness in the Smart City. In S. Vallor (ED.), The Oxford Handbook of Philosophy of Technology. Oxford University Press.

[25] Ethnographic research conducted in Amsterdam found that individuals have varying experiences with data collection based on their identity. Whether perceived or "real" these relationships will have an impact on how individuals navigate public space. Shazade Jameson, Christine Richter & Linnet Taylor (2019) People's strategies for perceived surveillance in Amsterdam Smart City, Urban Geography, 40:10, 1467-1484, DOI: 10.1080/02723638.2019.1614369.

informed of the sensor collection around them: their general location, the type of data, who is collecting, and for what purpose regardless of whether or not the data is personal in nature. With this information, those who use public space are better informed to self-govern. They may demand that unethical or simply unnecessary data collection be stopped or seek access to certain datasets. Public space is a shared entity that is difficult to avoid. If the consequences of people knowing data collection is occurring are greater than the consequences of the device itself, maybe the device does not need to be installed.

Data visibility in the end is a power concern.[26] Lack of public data awareness unjustly centers decision making power on those with the greatest economic and political control. Those left uninformed are unjustly distanced from a debate that affects them. Over the long term, ignoring the autonomy of members of the public will bias the ethical analysis of the IoT devices permitted to enter public space. By cowardly centering perceived consequentialist outcomes, digital public space distances itself from the public. It is plausible to see how this distancing could lead to a slippery slope of government and corporate surveillance under the argument that the end justifies the means. Contrastingly, if the public is welcomed to know and understand the sensors around them it is more likely that they will explore, ask questions, and hopefully also be able to demand that the IoT infrastructure benefits them directly.

[26] Algorithms themselves already hold power. If only a certain subset of the population gets to choose how, when, and where they are implemented then there will be an even greater imbalance. For more on the power of algorithms: David Beer (2017) The social power of algorithms, Information, Communication & Society, 20:1, 1-13, DOI: 10.1080/1369118X.2016.1216147.

The Challenges of Data Visibility in Practice

People value their privacy, even though they do not always behave that way. They tend to find it important, but hardly act upon it. This behavior is called the privacy paradox.[27] Generally, when people are left uninformed, they do not experience a violation of their privacy, because they have the expectation that they can move anonymously through public space.[28] Only when individuals are made aware, can they begin to understand the data collection and algorithms working in the background. Unfortunately, when "over-informed," privacy related information becomes overwhelming, causing individuals to start caring less about data collection and experience a loss of control.[29] This phenomenon is called privacy fatigue. At the same time, people feel disempowered when there is not enough transparency.[30] In practice, achieving data transparency ethically and effectively is challenging. This section will outline lessons from current legislation and initiatives, including the MX3D bridge in order to pinpoint needed regulations in the next section of the paper.

How people are informed about data collecting devices is dependent upon the interests of stakeholders with varying levels of power in the decision-making process. While this process should prioritize the general public, such as residents and tourists, often the relevant municipality or companies have more control over the design process. An example of this phenomenon was

---

[27] Maurits Martijn, Dimitri Tokmetzis, *Je hebt wél iets te verbergen* (5th edn, De Correspondent Uitgevers 2018).

[28] Martijn, Tokmetzis (n 28).

[29] Hanbyul Choi, Jonghwa Park, Yoonhyuk Jung, 'The role of privacy fatigue in online privacy behavior' (2017) 81 Computers in Human Behavior 42.

[30] 'People, Power and Technology' (*Doteveryone*, 2018) <https://doteveryone.org.uk/wp-content/uploads/2018/06/People-Power-and-Technology-Doteveryone-Digital-Attitudes-Report-2018.compressed.pdf> accessed 5 May 2021.

apparent at a public high school in China where the cameras went beyond the purpose of security. At this school, facial recognition was used to detect misbehavior such as being late to class and dating, which is not allowed between students.[31] While students knew smart cameras were present, they had no way of knowing the extent to which collected data was used to monitor their personal life. Since the students were uninformed they could do little to change their behavior or advocate for a different school policy.

In the absence of signage norms, transparency in public space will continue to be inconsistent and not geared towards the public's needs. Figure 5 shows an informationally dense sign designed for the MX3D bridge by van Stralen earlier in her research and Figure 6 displays a later draft from the company MX3D. Despite the differences in information shared, both signs at the time of creation appeared to be in line with current regulations. Additionally, both signs include a "sensor" symbol that had very low comprehension rates in survey tests.[32] There are numerous IoT infrastructures in Amsterdam currently collecting data with limited signage. Individual institutions might not want to provide more information than their counterparts thus leading to a race to the bottom in terms of data collection visibility.

[31] The school was one of X number of anonymous schools in the study Ken Anderson and others, 'A.I. Among us' [2019] Ethnographic Praxis in Industry Conference Proceedings 38.

[32] Survey respondents had the following assumptions about what the sensor icon meant: "Loud sound," "Risk of vibrations," "Watch out for sound? I don't understand it," and "Alarm signal." Four out of the thirty-two people who answered the question correctly guessed that the symbol meant that there were sensors present.

Figure 5. Information heavy sign (source: Authors)

Figure 6. Current MX3D draft sign (source: Authors)

Another concern with current transparency regulations is the specificity of data that they target. As shared in the introduction, GDPR only applies to personal data acquisition, which means anonymization allows bypassing informing the public. Similarly, academic research is given

special treatment under GDPR.[33] While this may be useful to not stunt research development, it is important that laws such as GDPR do not interfere with the rights of individuals to know about data collection in public space, especially when research stems from partially publicly funded research institutions. Other initiatives such as TADA and the Cities for Digital Rights are founded on values such as transparency and privacy, but do not give tangible regulations.

In the case of the MX3D 3D printed bridge it was unclear what the legal requirements were for such an IoT object. For example, the research group was originally looking for a local data storage opportunity preferably connected to or provided by the municipality of Amsterdam. Unfortunately, the city of Amsterdam was ill-equipped to host the data and agreed with the alternative offered by the consortium to store it with the Alan Turing Institute. This was decided partially so that the data would still be based in Europe (although not the European Union) and was deemed to be a better solution than hosting the data through Autodesk's Amazon Web Services servers. All of this decision making had to be done within the research group as it was difficult to find representation from the smart city of Amsterdam to offer assistance in the decision making processes.

The question of data visibility extends to data access. The smart bridge research group, for its efforts, desires a one year data hiatus in openly sharing the raw bridge data. Visualizations of processed bridge data should be available on the bridge website soon after the sensors begin running. After the first year concludes, raw data should be made accessible beyond the main smart bridge research group. Making data open is not as easy as it may at first sound. Given possible

[33] Miranda Mourby, Heather Gowans, Stergios Aidinlis, Hannah Smith, Jane Kaye, Governance of academic research data under the GDPR—lessons from the UK, *International Data Privacy Law*, Volume 9, Issue 3, August 2019, Pages 192–206, https://doi.org/10.1093/idpl/ipz010.

unforeseen consequences, it might not be ethical to expose all the sensor data to the public as it could turn the bridge into an experimental playground, which might disturb local residents and businesses. Anonymous data could also be paired with onsite investigation to target individuals. Moreover, those surveyed in the context of the bridge generally did not feel a strong right to have access to raw data.[34] While open data might be technically transparent, it fails in being translatable to the general population. Granting access to a lump of data without providing tools to comprehend the output also does not encourage individual autonomy. At the same time, the data could be useful to other researchers, who might make beneficial findings from publicly collected data.

The bridge is only one example in a growing sea of public IoT devices. Array of Things, a company that collects environmental data through "nodes," sensors with the words "Array of Things" on them that attach to light poles, proposes applications that might inform the public about the most populated route at night or the presence of ice further down the sidewalk.[35] Although nodes have been installed in Chicago and segments of the raw data are updated daily and available to the public,[36] it does not yet live up to its transparency potential because the barrier to inform oneself is too high. This leaves the average citizen unaware of the nature of the sensors and also unable to perceive any direct benefit. If one sees the sensor, which might be high up on

---

[34] Sage Cammers-Goodwin, Michael Nagenborg, 'From Footsteps to Data to Art: Seeing (through) a Bridge' (2020) special volume 8 Contemporary Aesthetics <https://contempaesthetics.org/2020/07/16/from-footsteps-to-data-to-art-seeing-through-a-bridge/> accessed 6 May 2021.

[35] University of Chicago, 'AoT is now an anchor partner in a new NSF-funded project called SAGE.' (*Array of things,* 2020) <https://arrayofthings.github.io/> accessed 21 April 2021.

[36] Array of Things, 'Five years, 100 nodes, and more to come' (*Medium*, 31 May 2018) <https://medium.com/array-of-things/five-years-100-nodes-and-more-to-come-d3802653db9f> accessed 21 April 2021.

a lamp post, and can read the text from far away, they might be able to find their way to the website. This threshold for data transparency can be lowered by publicly available data maps, where raw data is formatted and displayed in a more accessible manner.

Both Barcelona and Amsterdam offer web pages that show what sensors are located in their city.[37] Amsterdam does not provide open data access, but informs about the presence of different sensors and, where applicable, the privacy statement of the company responsible.[38] Starting October 2021 it will be mandatory to register all public sensors in Amsterdam.[39] Barcelona does share the collected data, but given that the map does not show any cameras throughout its whole city, it is likely that the sensor network is incomplete. Perhaps few people will find out about or reach out to this webpage. The threshold to learn about sensors in the city is lower, but the transparency is limited by lack of public data literacy and regulations for updating such sites.

One initiative that has extensively worked on making data visible to the public is Digital Trust for Places and Routines (DTPR). This open source communication standard originally developed out of Alphabet Inc's Sidewalk Labs as part of the Toronto Waterfront project before it

---

[37] Sentilo BCN (*Ajuntament de Barcelona*) <http://connecta.bcn.cat/connecta-catalog-web/component/map> accessed 21 April 2021; Sensors Crowd Monitoring System Amsterdam' (*City of Amsterdam)* <https://maps.amsterdam.nl/cmsa/?LANG=en> accessed 21 April 2021.

[38] Amsterdam does have some datasets open to the public which are not connected to their sensor map: Datasets (*Gemeente Amsterdam)* <https://data.amsterdam.nl/datasets/zoek/> accessed 6 May 2021.

[39] 'Inspraak Verordening meldingsplicht sensoren' (*Gemeente Amsterdam*, 24 February 2021) <https://bekendmakingen.amsterdam.nl/bekendmakingen/publicatie/inspraak/inspraak-sensoren/> accessed 23 June 2021.

prematurely shut down.[40] The standard later became fully independent in 2020.[41] They found that the public wants to be informed about the purpose of a technology, who is responsible for it, and an easy way to obtain more information.[42] They do this through hexagon shaped signs as this shape is less common in signage and therefore available for this new use case. Furthermore, hexagons can fit together nicely, so that familiar symbols may be combined to convey new meaning depending on the smart object. With the exception of a possible pilot in the city of Boston,[43] DTPR guidelines and iconography are currently not being put into practice.[44] The idea is promising, all design materials are open access and available for free use all over the world. However, their initiative offers such a variety of possibilities that it might be too overwhelming for companies and governing parties to start working with and it is likely that similar scenarios might end up having different ways of conveying the same message. Nonetheless, elements from this initiative can provide a strong basis for a new signage proposal.

When designing signage for the MX3D smart bridge, it was found through survey research that people would prefer to have limited information about a smart object within an urban space.[45]

---

[40] Sidewalk Labs is still under operation, but the Toronto project shut down May 7, 2020, before completion of the waterfront project; Jacqueline Lu, Chelsey Colbert, Patrick Keenan, 'How can we bring transparency to urban tech? These icons are a first step.' (*Sidewalk Labs,* 19 April 2019) <https://www.sidewalklabs.com/blog/how-can-we-bring-transparency-to-urban-tech-these-icons-are-a-first-step> accessed 21 April 2021.

[41] DTPR (n 5).

[42] Lu, Colbert, Keenan (n 41).

[43] 'Digital transparency in the public realm' (*City of Boston,* 8 October 2020*)* <https://www.boston.gov/departments/new-urban-mechanics/digital-transparency-public-realm> accessed 21 April 2021.

[44] DTPR (n 5).

[45] Naomi van Stralen, 'A Smarter Bridge' (Bachelor thesis, University of Twente 2020).

As also concluded by DTPR research, individuals are interested in the *purpose* of data collection and where to find more information.[46] A prior workshop on the bridge also found that people want to know who is responsible for the data collection.[47] This contrasts with what is generally put into practice, where the *type* of data collection may be specified, but the public is at best referred to an external site for more information as it is legally not mandatory to inform the public about data collection when it is anonymized according to the GDPR.[48] Legally, a website link or QR-code on location is sufficient when personal data is not anonymized, assuming that people will inform themselves.[49] This creates a higher barrier to entry to be informed and restricts those with limited digital literacy or without a smartphone.

As demonstrated by the examples in this section, there is a large diversity of IoT infrastructure and lack of clear guidelines towards data transparency. This lack of standardization makes it difficult to assign responsibility for data visibility and creates too much pressure and power on corporations to be ethical without providing guidelines on how to fulfill their duty. On the other side, civilians are expected to self-inform, navigating unstandardized signage and websites to comprehend data collection in public space. In order for data transparency to stop being an afterthought, regulations need to be put in place so that smart objects can converse with the public in a language all can understand. Recommendations for regulations are explored in the following section.

[46] Lu, Colbert, Keenan (n 41).

[47] Cammers-Goodwin, Nagenborg (n 35).

[48] General Data Protection Regulation (n 18).

[49] Point 58 of the introductory section of the General Data Protection Regulation states that any information addressed to the public could be provided in electronic form, such as a website.

PART 2. REGULATIONS FOR VISIBILITY

Even though there are international digital rights initiatives for IoT in public space, many citizens are still unaware of what a smart city is or how data is collected in the spaces they occupy. A gap exists between legislation and action. As described in the prior section, this might stem from the inconsistency of methods for showing data collection, diversity of IoT projects, and lack of a singular reporting system. This section will go more in detail about existing initiatives and what services need to be built to offer transparency and visibility. We make four recommendations that can be seen in Figure 7: standardized signaling, sensor registration, contextual data accessibility and increased data literacy education.

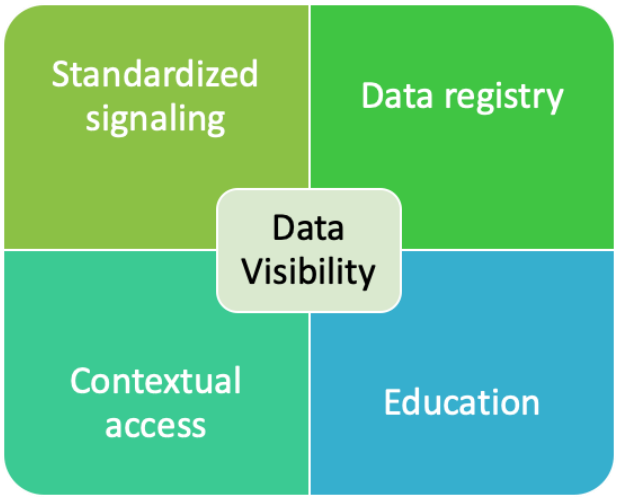

Figure 7. Four pronged approach to data visibility and transparency (source: Authors)

First we will address standardized signaling. It should not be possible to be unaware of public sensors if one is on the lookout for them, such a design would by definition be untransparent. It should be easy to connect from the IoT object to its purpose, function and owner. Visibility should not shift depending on the actors, but there should be indications of who the actors are, whether

they be governmental, academic, or commercial. Simple indicators such as color and familiar iconography may add clarity to public space without becoming too overwhelming. We believe that a DTPR based signage could fulfill all of these needs once simplified and internationally standardized in line with the Vienna Convention on Road Signs and Signals of 1968.[50] This convention is signed by 66 contracting parties in Europe, Africa, the Middle East, Asia and Latin America, and is the most universal guideline that is currently put into practice.[51]

Traffic signs offer a basic guideline on how to design for fast comprehension and visibility. Additionally, the sign shapes can be displayed through both analog and digital mediums. Consistency within the different forms of visibility such as on site, through an online government registry, or on an external website ensures that people are not overwhelmed by the presented information once they are educated about the language and icons. How people are informed influences perceptions and relationships with data.[52] If public data literacy increases, signage will not have to play as complete a role in informing the individuals. Currently however, they offer a first tangible step towards transparency and visibility

[50] United Nations Publications, 'Vienna Convention on Road Signs and Signals' (1968).

[51] '50 years on, the 1968 Conventions on Road Traffic and Road Signs and Signals are still at the core of road safety efforts worldwide' (*UNECE*, 7 November 2018) <https://unece.org/transport/press/50-years-1968-conventions-road-traffic-and-road-signs-and-signals-are-still-core> accessed 1 May 2021

[52] Cammers-Goodwin, Nagenborg (n 35).

Figure 8. Street sign inspired data collections signage (source: Authors)

By simplifying the DTPR proposal, it will be easier to achieve consistency, making it simpler for both those the producers of IoT objects as well as the public that has to interpret the signage. According to the MAYA principle, which stands for Most Advanced Yet Acceptable, something new should contain just enough elements for the user to understand it to make the new elements easy to adopt.[53] As shown in Figure 8, combining hexagonal signs will allow the viewer to be informed about the sensors present, whether they collect personal data, their purpose, who is collecting the data and where to find more detailed information. This allows the viewer to evaluate IoT infrastructure and form their own judgement. The public can gain autonomy by having the opportunity through awareness to opt-out or protest.

[53] Raymond Loewy, *Never Leave Well Enough Alone* (1st end, The Johns Hopkins University Press 1951).

Icons offer greater visibility to a larger audience compared to text, as they can be understood regardless of language or literacy.[54] By orienting the hexagon pointing upwards, it distinguishes itself from the octagonal signs. The colors, which are mostly primary colors, are adopted from existing traffic signs for greater familiarity and improved distinguishability for the visually impaired. In color theory, yellow is cheerful and represents optimism.[55] This can convey non-identifiable data, such as motion detection, or humidity. Blue signifies honesty and loyalty and can signify personal identifiable data. Red is left out, as this is a "warning" color and will come across as dangerous.[56] This could result in individuals feeling powerless towards their data being collected. Coincidentally, yellow and blue are also the colors used by DTPR. However, they use shades that deviate from traffic signs, which could make them less recognizable by the public.

The purpose of the data collection is positioned below the sensor data information, as well as the party responsible for the sensor and how to find more information. As these signs merely inform the viewer, they are white, in line with the Vienna Convention on Road Signs and Signals. Informative signs may either be dark with white or light-colored inscriptions or white or light-colors with dark inscriptions.[57] A website link is added to provide additional information in accordance with existing regulations such as the GDPR. As found through survey research as well

---

[54] Of the General Data Protection Regulation (2016) point 26 of the introductory section and Chapter III, Section 1, Article 12 point 7 that "information may be provided in combination with standardized icons in order to give in an easily visible, intelligible and clearly legible manner, a meaningful overview of the intended processing."

[55] Sean Adams, *The Designer's Dictionary of Colour* (Abrams, 2017).

[56] The Vienna Convention on Road Signs and Signals also states that red may only be used exceptionally on signs that have the purpose of informing. United Nations Publications (n 51) *Part I: Convention on Road Signs and Signals - Annex 1* (Section G, I. General characteristics and symbols, point 2).

[57] United Nations Publications (n 57) .

as by the Sidewalk Labs, there will be no added value to display the complete explanatory information at a sensor itself. Providing too much information on location may cause privacy fatigue and may be too overwhelming and irrelevant for most individuals. An example of how the proposed signage language can be put into practice is shown in Figure 9.

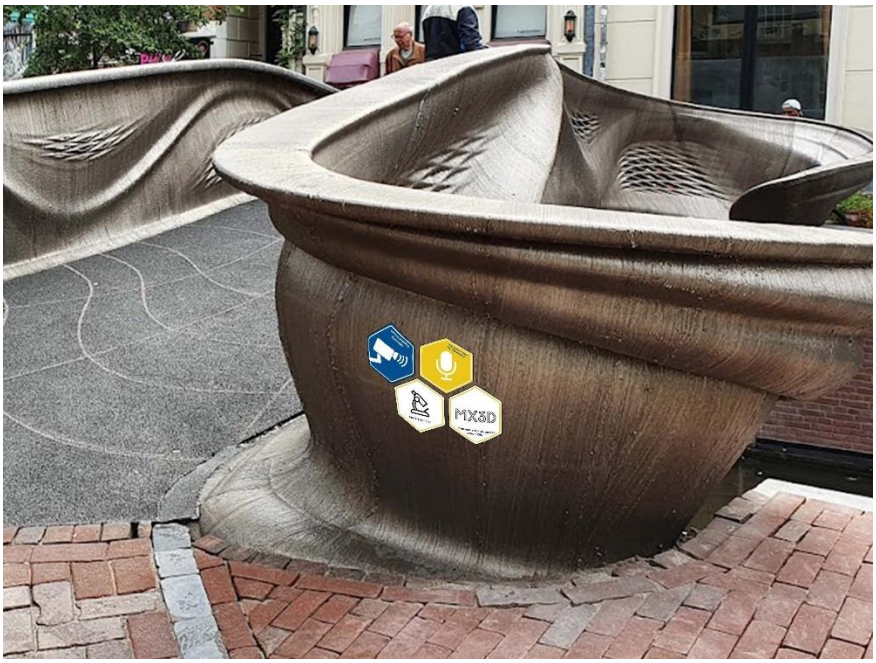

Figure 9. MX3D smart bridge signage prototype. Image, edited.[58]

Next we will cover the need for mandatory sensor registers. The logical follow up to standardizing physical transparency is to regulate the digital visibility of public data collecting devices. Currently, most signs for IoT infrastructure connect to disparate websites that have limited standards for what information should be shared. Ideally, public signs could link to the city's sensor map. This would allow the public to have one consistent platform and the option to

[58] The icons on the bridge provide a general example of the implementation of the proposed signage and is not a true representation of the sensors present on the bridge. Image bridge, edited. Kean Walmsley.

find out more about other sensors throughout the city. This might also be useful to those who want to plan their trips or the visually impaired who might miss signage. Mandatory public data collection registries, such as the one set to be enacted in Amsterdam fall 2021, bypass the issue of a company deciding how anonymous is anonymous enough to forgo the work of informing the public. The data registry system can be set up in a way that it forces all institutions to share the same information on the nature of their project so that public data collection information can actually be accessible to the public.

These data censuses must be open to the public and easy to access to give individuals agency to support, avoid, or fight against the presence of IoT technology in public space. Such tools make the public data acquisition truly a public issue. New changes should be highlighted so that it is easy to stay up to date with how the city is utilizing public sensors. Furthermore, there needs to be clear guidelines for public data operators to join, leave, and update the registry. Such a protocol would force sensor operators to have a clearer understanding of how they plan to use the sensor data. It is challenging to create transparency for the public when those installing the sensors are also unsure of how in the long-term they will use the data. Such a set up also ensures that the government stays aware of data collection in the city and distributes the burden of ethical decision making away from solely private actors.

Next we will discuss how to tackle open data via contextual access. Just because data collection is transparent does not mean the data itself is visible. Moreover, just because data is made available, does not imply that the system is transparent. Helen Nissenbaum's theory of contextual integrity suggests that people do want to share data, but they expect that data to only

travel within the appropriate domain.[59] For example, when one goes to the hospital they expect that the doctor has access to their medical history, but they might not want that information shared with their employer. In the public space context this might translate to security camera footage only to be viewed by the authorities and victims in the case of a suspected crime and not to be openly available to everyone. Contrastingly, data that actually does impact everyone, such as pollution data would make sense to expose publicly. Publicly generated data needs to be shared in a way that benefits the public as opposed to freeing information that only large institutions have the means to process and profit from analyzing. The *processed* data should in the end either be available or provide benefit to those moving through public space.

As with the prior two regulatory suggestions, without standardization, contextual open data might be challenging for the institution installing the IoT system. With standardization, however, the process could be made easier. Sensor registers could have built in APIs that allow data collectors to share relevant datasets, published research articles, and outcomes of the data collection. This makes the information easy for the public to find and for data collectors to share already processed information of value. Assuming that public data registers and appropriate signage are made mandatory it will be less likely that providing open raw data somewhere will be seen as a substitute for data collection transparency.

Finally, there is little point to increasing data visibility and transparency if people cannot understand enough to make educated decisions or if segments of the population are excluded. One way to target all groups is to start training data literacy at a young age through primary school education. Teaching all students through public school ensures that there is less chance of an

[59] Helen Nissenbaum, 'Privacy as Contextual Integrity' (2004) 79(1) Washington Law Review 119.

economic imbalance dividing the knowledgeable and uninformed about sensors in public space. This information may also reach parents through their children, thus educating the wider public about IoT data collection in public space. An educated public is also more likely to seek information because they are aware that the issue already exists. Education may also inspire students of varying backgrounds to think about how public space data might be useful for them or for them to invent their own solutions that might improve their community. This creates a knowledge powershift away from private industry and government and spreads it to a diverse younger generation. Additionally, to ensure that the older generations are not excluded, educational resources should be made available to the wider public to inform them on how to read data signage, where to find the data registry, and how to use contextually relevant open data sources. This information should be pushed through digital and print mediums once the standardization regulations and registries have been put into effect.

CONCLUSION

The MX3D bridge presents an example of how data transparency could be regulated to create cities where citizens have increased autonomy. Currently society is at a transitional stage where some initiatives are attempting to change the status quo and increase public awareness of data collection, but the visibility of IoT systems remains low, restricting individual autonomy. Decision making is still centered on those with the greatest economic and political control. The increasing amount of international regulations are guiding the world towards fairer data collection and more control over personal data. However, thus far, these regulations fail to generate greater data and algorithmic visibility in public space, especially for anonymous data collection. The gap between

rules and action needs to be narrowed. As with the case of the MX3D bridge, groups are stuck between informing the public in a way that increases the transparency and visibility on location or choosing a less informative method as there is no guideline on how visibility should be achieved.

Building and implementing the four pronged strategy suggested in this paper is likely to be expensive, but so is IoT infrastructure. Furthermore, the risks of not strategizing and implementing these tools now may have consequences for future citizens. Standardized signaling, a data registry, context based open data policy and data literacy education all work towards making data more visible, but cannot be achieved by independent initiatives alone. Governing parties need to work on the enforcement of data visibility to make data collection transparent to all. We need to create a fair environment where data collection is neither secretive nor threatening. At the current rate of development, technological advancements may exceed the capabilities of the current regulations in place to protect citizens. It is essential to develop practices that make public data collection a matter of public concern. Building regulations centered around individual data autonomy in public space may save heartache in the long run and force public IoT to actually be for the public good.

ACKNOWLEDGEMENTS

The research presented in this paper has been funded by the Netherlands Organisation for Scientific Research (project: Bridging Data in the built environment (BRIDE); project number: CISC.CC.018). Comments have been provided by Michael Nagenborg.