

Data commons in smart mobility – the road ahead?

Nathalie van Loon
Innovation Department
City of Amsterdam
Amstel 1, 1011 PN, Amsterdam, Netherlands
Nathalie.Loon@Amsterdam.nl

Rosalie Snijders
Faculty of Science, Information Studies
University of Amsterdam
Science Park 904, Postbus 94323, 1090 GH Amsterdam
Snijdersrosalie@hotmail.com

ABSTRACT

Mobility data collection and governance are mainly dominated by larger technology companies that gather all the data. Therefore, they also have exclusive control over what happens with the data. This calls for alternative data governance models. A viable alternative, introduced in recent years, is the data commons model. With this model, people can share their data on their own terms, while maintaining a certain amount of privacy. This model has been used with health data and scientific data, however, no viable example of a mobility data commons has thus far been found. This paper explores how local governments can facilitate a mobility data commons. And: is the commons a beckoning road for all of us?

KEYWORDS

Data governance, disruptive technologies, mobility data management, digital literacy, data commons, policy making.

1 INTRODUCTION

In the last decades, the concept of a smart city has grown in popularity both as a research subject and in government policies. Cities all over the world have started using technology to look for solutions that enable transportation linkages, mixed land uses, and high-quality urban services with long-term positive effects on the economy and sustainability of the city [1].

Smart cities are built on data. And one area where the generation and analysis of data have steadily increased is the mobility sector. App-based mobility services, like bike-sharing, scooter-sharing, peer-to-peer carsharing, and ride-hailing gather enormous amounts of information about how, when, and where people travel. And not only sharing apps, also other apps like weather apps or wayfinding apps generate data. Plus not only ‘smart solutions’ generate data but also ‘regular’ cars and bikes are becoming more and more mobile sensors in the city landscape by offering, to name just a few examples, ‘tracking services’ in case of theft, and cameras helping people to park.

In this context the City of Amsterdam aims to be a smart and mobile city, offering a large supply of mobility options; affordable, reliable, and accessible to everyone. However, most mobility data are enclosed by private companies, while the data

generated by these services can be of great public value. As the city of Amsterdam is also part of the ‘cities coalition for digital rights’ and aiming to be a number one city in the protection of its citizens digital rights, Amsterdam is looking for good examples in the governance of data and cocreation of public value together with citizens, local stakeholders and SMEs.

Considering the context, and considering the role of the municipality, this paper explores the following question: can or should a local government organize a data commons in order to enable parties to share data in a trusted, fair and economic way, while observing privacy and security concerns? This paper therefore shortly explores the ‘why’ and ‘how’ and evaluates the applicability of a data commons as a disruptive technology and framework. This paper is based on existing literature and interviews with experts from the municipality of Amsterdam and is structured as follows: section 2 will start with some background information to support the research question. In section 3 the concepts of a smart city and data commons are explored, and section 4 will present the conclusions.

2 BACKGROUND

In the last couple of years, data have become a valuable asset to our economy. Some have claimed that the world’s most valuable resource is no longer oil, but data [23, 49, 53]. A new form of capitalism has arisen where wealth is generated based on the accumulation, extraction, processing, and use of data.

The term Big Data has been on the rise since the start of the new millennium. Enabled by new and innovative technologies, companies can gather and analyse data from their customers or users and use it to their advantage. Digital data and information have become a critical economic, political, and social resource and most of this data is in the hands of just a few companies such as Amazon, Google, Facebook, and Apple [41, 43]. With this data, these few companies can have huge control and influence over human behaviour and societies. As a response, politicians, human rights movements and people in general have raised concerns about the misuse of their data. For many, it is not clear how much data these companies collect and what they do with it. As a result, people opt to not share anything with anyone and have started hoarding their data. However, data can be of great value for everyone if used in the right way. In the near future for instance, Artificial Intelligence will have to use data to play a role in the delivery of services [36]. If this data stay in the hands of big tech companies, the positive effects may never reach citizens.

As a digital rights city, therefore, it is of importance to look for new technologies that enhance public value and public benefit at the same time [43]. Citizens should have the power to decide on who they want to share their data with, under which rules, for

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
Information Society 2021, 4–8 October 2021, Ljubljana, Slovenia
© 2021 Copyright held by the owner/author(s).

what purpose and in a transparent manner. Data are (too) often regarded as a resource to be extracted for private profits, and technical developments have enabled technology firms to capture data from and about those who have not consented or have no viable alternatives. The view on data therefore must change from an asset that can offer a competitive advantage, to one of public infrastructure to ensure common welfare, which can be exchanged equally.

For this research, a smart city is defined as "A well defined geographical area, in which high technologies such as ICT, logistic, energy production, and so on, cooperate to create benefits for citizens in terms of wellbeing, inclusion and participation, environmental quality, intelligent development; it is governed by a well-defined pool of subjects, able to state the rules and policy for the city government and development". The city of Amsterdam has already become an example of how a smart city strategy can be implemented. With the above mentioned definition in mind, the main goal of smart cities is to improve the quality of life for its citizens in a sustainable way. At the same time, citizens also have the potential to be the main component of data acquisition. With the use of smartphones, the citizens can act as human sensors and help gather enormous amounts of data [29]. ICT can act as a platform to collect information and data to promote an improved understanding of how a smart city is functioning in terms of services, consumption, and lifestyle. Especially with mobility data, the input of citizens can be of great value [30].

While the potential of big data is explored on a daily basis in the development of new and possibly disruptive technologies, the potential societal disruption and ethical concerns attract less attention or even denial and/or apathy. This while multiple studies show that, with the creation of intelligent mobility systems in smart cities, the potential for intrusive surveillance is increased [31] and that the types of data used are privacy-sensitive [32]. Location history data, for instance, can act as an identifier of its users [33, 34]. Also bias in data can be a multiplier of societal injustice, as the Dutch 'toeslagenaffaire' [35] has shown, framing approximately 26.000 parents as possible fraudsters, based on their (second) nationality. Also multiple organizations may have multiple policies and rules regarding the protection of the data of their users. However, this is not always transparent - while it may lay in everyone's interest to share this data [36]. Therefore, one of the main challenges of the use of big data are privacy, transparency, and bias.

3 Data Commons

There are various definitions in use for commons and also for data commons. In general, the Nobel prize winning work on commons by Elinor Ostrom in 1990 is used as a reference for any such definition. Ostrom successfully described the commons as a governance model rather than open access to resources and introduced the commons as a framework to value various historical and contemporary social movements. In short one can define the commons as a commonly owned and managed (common pool) resource. More elaborate, Ostrom identified 8 design principles of stable common pool resource management in her groundbreaking work 'Governing the commons. The evolution of institutions for collective action.' [3, 6, 18, 19, 27].

3.1 Design principles data commons

Principles can be described as general rules and guidelines which a system architecture must follow to be as productive and cost effective as possible. Principles help guide the use and deployment of an architecture. Also principles may help identify concerns stakeholders might have that a system can address. Each principle should have a rationale and implication associated with it. This can help with promoting the acceptance and understanding of the principles [10, 25]. Here, we adapted and 'translated' 7 of Ostroms 8 design principles - in a first attempt - to rationales and implications for data commons [5].

- (1) Define clear group boundaries:
 - *Rationale*: Who can use the data should be clearly defined and should be easily identifiable
 - *Implication*: An individual using the commons may require identifying information before allowing access to the commons. Additionally, the data sets should be easily identifiable. With this in place, poaching can be easily detected [23].
- (2) Match rules governing the use of common goods to local needs and conditions:
 - *Rationale*: The rules of governing the data commons should be matched to the local needs of the users. Since no data commons and its environment are the same.
 - *Implication*: Setting up the rules and guidelines of the use of the commons should include the local users of the commons. Therefore, citizen participation is a crucial part of a successful commons.
- (3) Ensure that those affected by the rules can participate in modifying the rules:
 - *Rationale*: Both the data producer and user should be able to benefit from the data commons and be protected.
 - *Implication*: All parties within a data commons should be able to change the conditions of the data commons, with agreement from all parties. The use and production in the data commons should always be in balance.
- (4) Make sure the rule-making rights of community members are respected by outside authorities:
 - *Rationale*: The rules and regulations of the commons should be respected by the local authorities, must be recognized as legitimate by the authorities.
 - *Implication*: Local authorities shouldn't be able to change the rules without the consent of the parties involved.
- (5) Develop a system, carried out by community members, for monitoring members' behaviour:
 - *Rationale*: Monitoring of the data commons is needed to ensure that the data is used fairly.
 - *Implication*: Unauthorized use of the data should be detected. In the case of a data commons, this could be a moderator, since the commons are not in a physical place. Ideally, this is done by the user community.
- (6) Use graduated sanctions for rule violators:
 - *Rationale*: Users and producers in the data commons that violate its rules should not be banned directly.
 - *Implication*: A gradual system needs to be set up.
- (7) Provide low-cost accessible means for dispute resolution:
 - *Rationale*: When issues within the commons come up, the dispute would have to be resolved in an informal, cheap, and straightforward manner. This way problems are resolved, rather than ignored
 - *Implication*: A process for conflict resolution should be created that is perceived as fair by all users of the data

commons. A mechanism for rule enforcement and for dealing with violators needs to be set up and discussed by all involved parties.

Concerns

The incorporation of the above-mentioned design principles can be a measure of success when organizing a data commons. But can they also be used to address the concerns the relevant stakeholders might have?

3.2 Citizen participation

Since citizen participation is a necessary step when organizing a datacommons and is essential for two design principles of a successful data commons, a major concern when it comes to a local government organizing or facilitating a datacommons is the participation of citizens. Is this a ‘*contradictio in terminis*’ or can and should the government act as a facilitator or incubator? Looking at the participation ladder by Arnstein [2] there is, indeed, a world to win, also calling for a different role of the government: a ‘*co-creating government*’ or ‘*co-city*’.



Figure 1: Levels of participation

Transparency

Another important concern is transparency; in order to achieve a successful mobility data commons, the municipality needs to be transparent about every part of the data commons. To achieve full transparency, openness of all operations within the data commons is required, so that citizens if needed, can hold the consumers of the data accountable and are allowed to withdraw their consent [24]. However, measuring transparency within a data commons can be a tricky task. The question is not only how much information is available and under which terms, but is also a question of equality in the accessibility and usability of that information. Transparency is increased when the data within a data commons is given a proper context and, therefore, its users can use and understand the data without confusion. Transparency should cover all of these aspects of data access: physical access, intellectual access, and social access [13]. In the case of a data commons, physical access can refer to the ability to reach the content of the commons, social access is the ability to share the content of the commons and intellectual access is the ability to fully comprehend the content [7, 4], sometimes also referred to as ‘*digital literacy*’. Not only in Amsterdam, but in more cities in the digital rights coalition, the Covid-19 pandemic

and subsequent lockdowns showed that a lot of families don’t have access to technology when public services like libraries and schools are closed. And how can Amsterdam residents take ownership over their data if they don’t have access to technology, know where to access their data or how to object to their data being used? By introducing a ‘*digital agenda*’ [41] the city of Amsterdam is working on overcoming this divide and promoting and protecting digital rights, yet agency is complex and scattered. Also, the use of data and which algorithms are used should always be disclosed to the contributors of the data. Amsterdam has made a first step by introducing an ‘*Algorithm register*’ [16]. But can a commons be organized in such a way that no one has access to a contributor’s data without their permission?

Monitoring and validating

This also raises the question if local governments can organise the monitoring of the use and validation of data. A solution could be implementing an interoperable context-aware meta-database architecture [15]. This type of architecture is context-aware and allows permissions and policies to be attached to the data. Additionally, due to its flexibility, trust norms can be changed and can account for increased transparency and accountability. This is an architecture that associates data with user permissions and policies which enables any consumer to handle the data in a way that is consistent with a contributor’s wishes [21]. This is a method that could increase accountability in a decentralized data ecosystem like a data commons. However, this method does thus far not provide a way for community members to contribute to monitoring the behavior within the community.

Concerns

Interoperability is a practical, yet very prominent concern when organizing data commons [7, 8] since a data commons is not only about access to data, it is also a platform for data experimentation and interaction. Technically, a data commons is a repository of personal manifests that describes the access and usage rights of all data generated by an individual within a digital service. Therefore, the data commons must regulate relationships between the organizations and individuals that use and share ownership of the data. This way, data commons help citizens having a say in what data they want to share and under which conditions. Also data commons could provide users easy access to their own data, information about who has access to their data and what they could do with this information. However, for this to be successful also trust needs to be built between the different parties participating in the commons. As our last concern we raise the question on the definition and the narrative. The commons, although part of an important and impactful historical movement, that, amongst others, created the guilds in the Middle Ages, the common land movement in the UK and, more recently, knowledge commons Wikipedia [11], mutuals like ‘*broodfondsen*’ in the Netherlands and citizen energy communities in most European countries, are not part of our current, dominant, narrative. Although the European Union and Dutch government have legal frameworks in place for several types of commons - in housing and energy for instance- no real understanding of the potential public value or even clear definition of a data commons currently exists.

4 CONCLUSION

For now, the larger technology companies dominate the data collection in the area of mobility. As a result, these companies have exclusive control over what happens with the data the citizens of a city generate. In this paper we described how this ‘enclosure’ of data by big tech builds a powerful value driven case for cocreating and/or facilitating commons in mobility data as a local government. Although a clear pathway on how to organize a mobility data commons is not yet available, the road ahead is one of cooperation, building trust between participants and experiment. By taking it one step at a time, setting clear boundaries and rules that are understood by partners involved and, obviously, involving citizens in every step. However, considering digital literacy and other possible constraints for citizen participation, careful thought on how to involve citizens -for a longer period- is paramount. One suggestion would be to just ‘follow the music’: there is a vibrant movement of active citizens communities and SMEs in town, how can the local government cooperate towards the creation of a data commons in mobility as a spill-over effect from these efforts? This way data commons can prove to be an alternative for apathy and distrust in big tech, contributing to a strong and growing narrative on local cooperation.

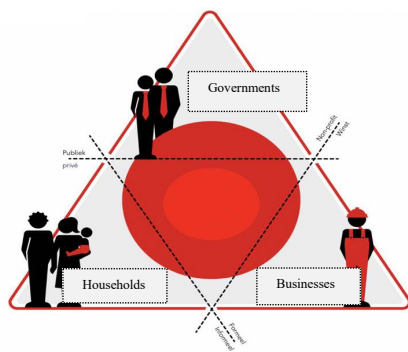


Figure 2: the third sector model [37]

ACKNOWLEDGMENTS

This paper is partly based on interviews with public servants and experts at the city of Amsterdam, a research done by Rosalie Snijders in Q2 of 2020. During this period Rosalie was an intern at the city of Amsterdam, writing her master thesis, supervised by Nathalie van Loon, working at the city of Amsterdam and Leon Gommans and Frank Nack, both working at the University of Amsterdam. This paper is an adaptation by Nathalie van Loon, written in the context of the Urbanite project: <https://urbanite-project.eu/>, under grant agreement #870338.

REFERENCES

- [1] V. Albino, U. Berardi, and R. Maria Dangelico. Smart cities: Definitions, dimensions, performance, and initiatives. *Journal of Urban Technology*, 2015.
- [2] S. R. Arnstein. A Ladder Of Citizen Participation. *Journal of the American Planning Association*, 1969.
- [3] B. J Birkinbine. Commons praxis: Toward a critical political economy of the digital commons. *TripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society*, 16(1):290–305, 2018.
- [4] G Burnett, PT Jaeger, and KM Thompson. The social aspects of information access: The viewpoint of normative theory of information behavior. *Library & Information Science Research*, 30(1):56–66, 2008.

- [5] https://en.wikipedia.org/wiki/Elinor_Ostrom
- [6] J.B. Fisher and L. Fortmann. Governing the data commons: Policy, practice, and the advancement of science. *Information & Management*, 47(4):237 – 245, 2010.
- [7] R. L. Grossman, A. Heath, M. Murphy, M. Patterson, and W. Wells. A case for data commons: Toward data science as a service. *Computing in Science Engineering*, 18(5):10–20, 2016.
- [8] R. Grossman. A proposed end-to-end principle for data commons, 2018 (accessed April 5, 2020).
- [9] https://assets.amsterdam.nl/publish/pages/964754/agenda_digitale_stad_tussen_apportage_2019_-_2020.pdf
- [10] Van Haren. TOGAF Version 9.1. Van Haren Publishing, 10th edition, 2011.
- [11] I. Abaker T. Hashem, V. Chang, N. Badrul Anuar, K. Adewole, I. Yaqoob, A. Gani, E. Ahmed, and H. Chiroma. The role of big data in smart city. *International Journal of Information Management*, 36(5):748 – 758, 2016.
- [12] C.Hess and E. Ostrom. Introduction: an overview of the knowledge commons. *Understanding knowledge as a commons: from theory to practice.*, 2006.
- [13] C. Humby. Data is the new oil. *Proc. ANA Sr. Marketer’s Summit*. Evanston, IL, USA, 2006.
- [14] P.T. Jaeger and J. Carlo Bertot. Transparency and technological change: Ensuring equal and sustained public access to government information. *Government Information Quarterly*, 27(4):371 – 376, 2010. Special Issue: Open/Transparent Government.
- [15] P. T. Jaeger and G. Burnett. *Information worlds: Social context, technology, and information behavior in the age of the Internet*. Routledge, 2010.
- [16] S. Maguire, J. Friedberg, M. H. Carolyn Nguyen, P. Haynes. A metadata-based architecture for user-centered data accountability. *Electronic Markets*, 2015.
- [17] <https://ai-regulation.com/amsterdam-and-helsinki-launch-algorithm-and-ai-register/>
- [18] H. Mehr. *Artificial Intelligence for Citizen Services and Government*. Harvard Ash Center Technology & Democracy, 2017.
- [19] E. Ostrom. *Governing the commons: The evolution of institutions for collective action*. Cambridge university press, 1990.
- [20] E. Ostrom, R. Gardner, James Walker, J. M Walker, J. Walker. *Rules, games, and common-pool resources*. University of Michigan Press, 1994.
- [21] F. Pasquale. From territorial to functional sovereignty: The case of amazon. *Law and Political Economy*, 6, 2017.
- [22] E. Politou, Efthimios Alepis, and Constantinos Patsakis. Forgetting personal data and revoking consent under the gdpr: Challenges and proposed solutions. *Journal of Cybersecurity*, 4(1):tyy001, 2018.
- [23] B. Prainsack. Logged out: Ownership, exclusion and public value in the digital data and information commons. *Big Data and Society*, 2019.
- [24] N. Purtova. *Health Data for Common Good: Defining the Boundaries and Social Dilemmas of Data Commons*, pages 177–210. Springer International Publishing, Cham, 2017.
- [25] S. Ranchordás. Nudging citizens through technology in smart cities. *International Review of Law, Computers and Technology*, 2019.
- [26] E. Schlager. Common-pool resource theory. *Environmental governance reconsidered: challenges, choices, and opportunities*, pages 145–175, 2004.
- [27] S. Spiekermann, Alessandro Acquisti, Rainer Böhme, and Kai-Lung Hui. The challenges of personal data markets and privacy. *Electronic markets*, 25(2):161–167, 2015.
- [28] J. Yakowitz. Tragedy of the data commons. *Harvard Journal of Law & Technology*, 25(1):1, 2011.
- [29] M. Srivastava, T. Abdelzaher, and B. Szymanski. Human-centric sensing. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 370(1958):176–197, 2012.
- [30] J. Vázquez Salceda, S. Álvarez Napagao, J. A. Tejada Gómez, L. Javier Oliva, D. Garcia Gasulla, I. Gómez Sebastià, and V. Codina Busquet. Making smart cities smarter using artificial intelligence techniques for smarter mobility. In *SMARTGREENS 2014: proceedings of the 3rd International Conference on Smart Grids and Green IT Systems*, pages IS7–IS11. SciTePress, 2014.
- [31] M. Büscher, P. Coulton, C. Efstratiou, H. Gellersen, D. Hemment, R. Mehmood, and D. Sangiorgi. Intelligent mobility systems: Some socio-technical challenges and opportunities. In *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering*, 2009.
- [32] Y. Alexandre De Montjoye, C. A. Hidalgo, M. Verleysen, and V. D. Blondel. *Unique in the Crowd: The privacy bounds of human mobility*. Scientific Reports, 2013
- [33] C. Bettini, X. S. Wang, and S. Jajodia. Protecting privacy against location-based personal identification. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2005.
- [34] Hui Zang and Jean Bolot. Anonymization of location data does not work: A large-scale measurement study. In *Proceedings of the 17th Annual International Conference on Mobile Computing and Networking, MobiCom ’11*, page 145–156, New York, NY, USA, 2011. ACM
- [35] <https://nl.wikipedia.org/wiki/Toeslagenaffaire>
- [36] Lina Zhou, Shimei Pan, Jianwu Wang, and Athanasios V Vasilakos. Machine learning on big data: Opportunities and challenges. *Neurocomputing*, 237:350–361, 2017.
- [37] V. Pestoff, Third sector and cooperative services -An alternative to privatization. *Journal of consumer policy* 15(1): 21-45, 1992.