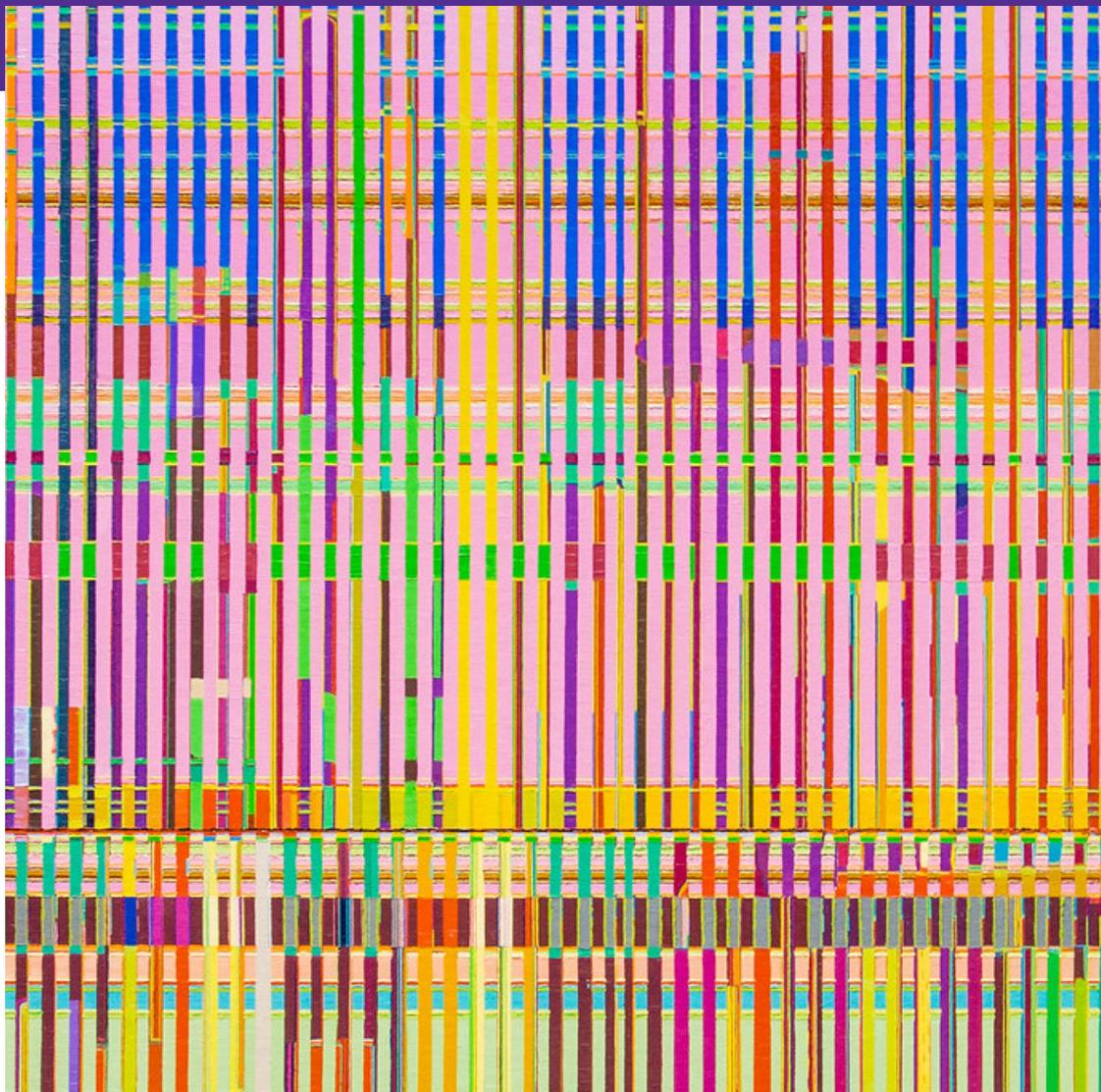*Socrates Schouten*

A strategy
for urban data

*How to develop collaborative data projects – for citizens,
urban innovators, researchers and policy makers*

AMS
AMSTERDAM INSTITUTE FOR
ADVANCED METROPOLITAN SOLUTIONS

waag
technology & society

DIGITAL
SOCIAL
INNOVATION

Table of contents

Preface

Many cities in Europe face major societal challenges in the areas of energy transition, climate adaptation, food security, mobility and social welfare. Solutions to these challenges are often so complex that collaboration is required between many parties from different domains (physical, spatial and social) and across various scales. In response, the city is increasingly conceived as a *living lab*: an environment where citizens and stakeholders work together on smart and creative solutions to urgent problems. And increasingly data play pivotal role in this as a means to inform, facilitate and enable interactions taking place in the city. It is in this regard that the idea of '*data commons'* is being brought to the fore, suggesting inclusive production of, and access to, data by a community that includes practitioners, researchers and citizens.

What exactly are data commons? What kind of applications do they appear in? And who are the parties involved in the research and development of successful urban solutions? In this report, Waag and AMS Institute investigate how new design methods and data technologies can be optimally used for research and experimentation. Through literature review, interviews with stakeholders and public sessions about 'Designing the City' we explored what possibilities there are for the sharing and reuse of data from projects executed in the urban context. Sharing data, if done wisely, has tremendous potential for research, innovation, new business, a better organised society and more civic involvement. Reuse of data promises large improvements in the time and pace of (scientific and civic) discovery. We formulate lessons learned and recommendations to guide these processes and anticipate the associated challenges. These lessons are especially geared toward 'intermediary' organisations and 'system actors' that want to support civil society with the question of shared data management. But first we need to make sure we are asking the right questions—to the right people.

This report was written in the context of the development AMS Institute's knowledge portal (knowledge.ams-institute.org, currently early stage) and the DSI4EU project (digitalsocial.eu).

We wish you an inspiring reading.

Introduction

Here is the old pitch: a citizen (let's call her Alice) has a good idea and is now looking to realise it. Therefore Alice needs a platform with lots of open data and a marketplace of ideas and challenges. This will give her everything to get started and make the city a better place. In short, Alice simply needs a data portal where she finds what she needs.

In practice, Alice stumbles and falls. She gets lost on a platform that contains little more than the *promise* of useful information—if she finds the platform at all. Data literacy, anybody?

Fans of Alice know that Bob is never very far away. This Bob is a social designer. He knows very little about data and technology, but he knows how to organise a process. Alice and Bob feel that their need for information is probably a simple thing to address. And they are right: Setting up a data portal is not the challenge; making it accessible and valuable is.

So here is the new pitch. We imagine a city that boasts initiatives of many different natures and sizes. All of these initiatives have an *open process orientation,* so they can, at intervals at least, absorb new ideas and people. Do you need data to succeed? There are organisations and platforms that can help you with your data questions, but do so by offering a routine that considers all adjacent issues just as intimately.

Storing data is simple—or maybe it is *complicated*, but it is never *complex*.[1] The larger context *is* complex, however. Any collaborative project involving data is set in a specific dynamic composed of different actors, laws and rules, expectations, levels of expertise, incentives and relations. The trick is to curate an accessible process, allowing citizens, communities and (say) NGOs to jump on the bandwagon when they have something to ask, to contribute, or both. This helps to navigate and manage a context that can be complex, or chaotic, indeed. The important questions to ask are: what are the initial and potential communities whose needs we are about to address? Given that community, to which levels of specialisation and creativity should we be open, and what type of language can we afford to use? What degree of complexity do we assume to apply and are we ok with that? And given that context, who should take the lead? What degree of institutionalisation—i.e. guidance by large, established institutes—is desired or required?

This report advocates the creation of **data commons**: data collections that are maintained and managed by communities of citizens in collaboration with other (local) parties. Data commons combine technical and organisational solutions so that citizens can exercise maximum control and there is shared value creation. The data commons approach favours open-ended data projects and seeks to avoid strongly hierarchical, tech-driven solutions. In the commons, investing in an engaged and skilled user community is key. At the same time, commons approaches keep an eye on the needs of the larger public and seek to involve tho-

---

1   *Simple, complicated, complex* and *chaotic* are the four main conditions of systems, be they ecosystems, financial markets or automobiles. *Simple* and *complicated* systems are predictable, but vary in their dimensionality. *Chaotic* and *complex* systems are both (largely) unpredictable, but vary in the meaningfulness of their behaviour.

se that can help sustain the resource. This report aims to substantiate the need for a commons approach for data.

In chapter **one**, we distinguish six hypothetical levels of complexity that demonstrate the different needs for institutional support and different types of stakeholders and collaborations involved. In chapter **two**, we discuss the state of the art of co-creation, living labs and collaborative data management. It also reports on our findings about the data needs of practitioners and citizens, and summarises the different insights into a suggested data process flow, which borrows from more and less familiar processes. Finally, chapter **three** provides recommendations for organisations that want to support civil society with their shared data management.

Waag 2016.

1. Six levels of data collaboration

No two data sharing initiatives are exactly alike. The nature of the data, the specialisation, skill and involvement of the participants and the continuity across projects may differ widely. The findings and recommendations presented in this report will not be applicable in every situation where data collaboration is sought.

Below we present six use cases that range across six (hypothetical) levels of complexity. These levels have been conceptualised based on discussions that took place during multiple workshops held in 2018 and 2019 in the context of the DSI4EU project and a shared project of Waag and AMS Institute. They describe six different situations, which look for roughly the same outcomes, but do so in disparate contexts. Springing from the same set of 'data sharing' ambitions, we found that we were having quite different conversations and were arriving on different conclusions. That is why we limit the scope to a narrower set of use cases, even if the broaders lessons may apply across the board.

These six levels mainly aim to demonstrate the different needs for institutional support depending on what questions are being asked and whom is thought to be the user, or beneficiary, of the data sharing effort. This report focuses on the middle level: a scenery in which small-scale organisations are the beneficiaries, and, often, initiators of data solutions organised in common. This level is situated between the 'citizen' level (where ordinary citizens are beneficiaries) and the 'institutional' level (in which case the game is played largely among well-established organisations). The middle level selected here does not live in isolation: collaboration with institutional partners is the rule rather than the exception, and they consistently aim to bring impact for citizens. But the driving logic is significantly different than that of the projects in other strata.

What becomes clear of the table below is that there is quite some variance in requirements, which suggests different design approaches. Some needs, certainly in the first three levels, revolve around finding relevant people and ideas (collaborators, target group, customers, certain knowledge). This question is more *incidental* than structural in nature. The higher-level examples (four and above) are more *structural*, aiming to develop a (bespoke) collaborative platform, enable new services, and be able to obtain a higher degree of (system level) knowledge. The expertise level and the type of economy that drives it will likely also be quite different.

As said, this report focuses on the intermediate level of civil society organisations, social entrepreneurs, SMEs and 'city makers'—semi-professional innovators with the common good at heart. The recommendations are mostly directed at system level actors such as governments and knowledge institutions. As an example, the next section discusses how the need for data collaboration is felt in the food sector.

Table 1. *The nature of collaborative data projects in six levels. This report focuses on the middle range where civil society organisations (CSOs) and small and medium enterprises (SMEs) are the main actors and beneficiaries (n° 3 and 4).*

Level	Example	Main stakeholders involved	
1	*Individual looking for practical information/insight.* A citizen wants to collect green waste for a composting project. How to find enthusiasts to collaborate with? What research has been done on voluntary composting?	Citizen: user Public sector: data manager	Citizen focus
2	*Local initiative occasionally requesting data.* Stadsdorp Nieuwmarkt, a group of residents, wants to improve social ties in the neighborhood to combat loneliness and address basic care questions. How to identify the people that need it? How to involve the non digitally savvy?	Group of citizens: user Public sector: data maintainer	
3	*Local initiative seeking to grow and scale.* A food cooperative wants data on the sustainability and healthiness of the conventional food system as a baseline to compare their impact with. They would also like to cooperate with other practitioners in the city.	Local initiatives, organised 'peer to peer', supported by (public) institutions	CSO / SME focus
4	*Tailor-made data exchange for SMEs and social entrepreneurs.* The city wants to organise bicycle data at city level to allow local sharing concepts. This can help the city and its residents to co-design and achieve policy objectives.	Quadruple helix, focused on local actors.	
5	*Systematic question, city level.* A city wants to better organise health sector data, to coordinate supply and demand for in-house care on the neighbourhood level.	Triple helix. Data also made accessible for civic stakeholders.	Institutional focus
6	*Systematic question, regional level and higher.* Comprehensive ambition to build a digitised circular economy and energy market based on real time asset-level data.	Triple helix. Data also made accessible for civic stakeholders.	

Case study: food system data

This section discusses, by way of example, one context in which we can find both the challenges and the opportunities of collaborative data projects on the level identified as relevant for this report: the food system.

What is at stake? On both ends of the production chain, food has a large impact: on the field it leaves a large environmental footprint; in cities, the consequences of logistics and consumption are felt. Starting with the first, agriculture accounts for 40 percent of European land use, and 44 percent of European water use; soil quality across Europe is deteriorating; and animal farming contributes significantly to greenhouse gas emissions. At the other end of the chain, one third of food bought is wasted. And food has a major impact on the city. Supplying shops with products, and increasingly the delivery of meals and food boxes to consumers, makes up a large part of urban traffic, and hence noise and air pollution.

A growing number of cities are moving towards regional food policies, with a focus on healthy nutrition, resilient supply and innovating urban agriculture. The 2015 Milan Urban Food Policy Pact drove many cities to adopt urban food policies—with both mayors and civic networks taking the initiative. Meanwhile, the innovative power of ICT is visible in the development of solutions that support the relocalisation of food. Often, multi-stakeholder platforms cooperate to design new policies which incorporate data solutions. The Dutch *Taskforce Korte Keten* ('Taskforce Short Chain') for example is an initiative of Dutch parties who want to take advantage of the opportunities for regional agriculture. This coalition wants to improve the connection between citizens and farmers through IT technology, community formation, social entrepreneurship and citizens' participation. According to the Taskforce, shorter food chains can stimulate demand-driven production, creating more opportunities for sustainable agriculture and the closing cycles up to a healthier food supply in the city. Similarly, the platform *Voedsel Verbindt* ('Food Connects') addresses food challenges in the Amsterdam Metropolitan area. Data and knowledge sharing offer opportunities to make food provision smarter and healthier, but there are also many challenges for the various parties.

Examples and interviews were gathered during workshop sessions that were conducted by Waag and AMS Institute in early 2019. The workshops were organised based on the premise that the knowledge produced both by researchers and by urban food projects can be of great value once shared, but sharing data is not straightforward at all. In addition, we interviewed Jessica van Bossum, researcher at the Amsterdam University for Applied Sciences (HvA) and in that capacity involved in *Voedsel Verbindt*, who has been taking steps towards a 'basic data set' for food in the city (see page 10). The aim of that data set should be to offer a common facility to all those that want to make analyses, make evidence-based decisions or innovate solutions. Our second interviewee, Anne Bruinsma of Farmhack (see page 11), is active on the other end of the food chain: farms and farmers. Farmhack's approach is to innovate using the data that is available from tractors, analytical instruments, official censuses and citizen scientists.

As becomes clear both from the 'basic data set' case and the experience of Farm-

hack, it is not that easy to make the solution work for the end-user (which could be farmers and consumers alike). For digital technology and data approaches to succeed in 'the field', depends not just on the product's durability and its service level, but also in proving your worth in a more general sense. The application should be easily available and succeed in different (sometimes challenging or unexpected) use cases. As a further complication, most technology functions as part of a software (and hardware) 'ecosystem' rather than isolated, which exponentially increases the effort of delivering a well functioning solution.

In our workshops, there was a general acknowledgement that projects aimed at sustainable and local food production should include the activation of citizens. The big aim is to turn the 'regular consumer' into somebody who is engaged, skilled and in control. People could be share-holder or co-owner of a solution, engage in self-production, or be active in neighbourhood projects. But in the workshops the question arose what stepping stones could bring us towards such an ideal situation. During the workshops, focussing on the process of data generation and sharing rather than the technology of data storage was identified as key enabler for better knowledge sharing among practitioners and system actors.

The next chapter will provide the argument for middle-level data collaboration. It describes why the need for collaborative projects is felt today more strongly than ever, given the challenges we face as a society. We also know that data is increasingly central in addressing challenges and realising collaboration between researchers, problem owners and urban stakeholders. That is why we need to build on current lessons about co-creation and data collaboration.

Ilyas bin Sarib, Waag 2019.

Interview: Jessica van Bossum, Amsterdam University of Applied Sciences (HvA)

Jessica's research work revolves around 'data-driven food system design'. At the Amsterdam University of Applied Sciences she created a map of actors in the food system using public data. Her aim is to make data available to researchers and social entrepreneurs working towards a healthier and more sustainable food system. That is also her aim as coordinator of the working group 'data & citizen science' of the newly launched 'Voedsel Verbindt' platform.

What are the objectives of the data working group?

To create insight into which data really matters. Where could we obtain useful data streams and how we can make this data accessible (up-to-date, with quality control, with the consent of data owners, published with context, and with rules to prevent misuse / abuse). In particular I'm researching what *basic data set* everyone should have at their disposal to allow a critical mass of people to work on a healthier and more sustainable food system.

What are your take-aways from the workshop?

I got a clearer view of use cases and the potential relevance of the basic data set. Such as using data as evidence base for new initiatives, for the validation of policy measures, and to develop business cases and innovations (e.g. an app to inform consumers about sustainable food choices, or the rating of entrepreneurs for sustainability). I also learned that the 'pitch' for making a basic food data set is not rock solid. But if a data resource is a public good, as in funded by public money, the relevance for society must be clear. After a while we arrived at the image of a 'CBS [Netherlands statistical agency] of the food system', which is quite appealing. Apart from the prospect of a full-fledged open data repository, they set a standard in terms of quality control, granting access at different levels for different users, making sure that the public can use the data, and developing a good revenue model.

Interview: Anne Bruinsma, FarmHack

Farmers are generating more and more data, but currently the possibilities of data and technology in agriculture are underutilised. There are many societal challenges that can be translated into innovation opportunities for the farmer through the smart use of data and technology. The mission of Farmhack is to build a rich and diverse ecosystem around the farmer, consisting of coders, hackers, developers, planners, designers, domain experts and officials.

What are the objectives of your project?

FarmHack supports farmers to do well in the digital age. There is a big wave of investment in Agritech and smart agriculture, but many farmers are not prepared or aware of opportunities and pitfalls. We use instruments like hackathons to connect the tech and agriculture world. Besides focus on smart agriculture, we also focus on smart provision of public services by farmers, such as care for the environment. We aim to develop knowledge, data, yardsticks, soft and hardware. All in all, there is a tremendous data resource in farming but it is not used optimally, and the benefits farmers get out if it is quite skewed compared to what big Agritech gains by all of this data.

What are your insights about data sharing?

We are all for building custom data sharing platforms, managed by communities of farmers or by social enterprises, but it is also a huge effort. That is why we embrace industry-driven initiatives just as much—but under the right conditions, including a fair share for the farmer. We think we can achieve that. Industry platforms often work very efficiently, and with their APIs allow you to add your own modules. I'm thinking of an initiative like JoinData, run by the likes of FrieslandCampina and AgriFirm. That said, we need to advocate for more focus on optimising for the end-user, the farmer. Is there a possibility to use the investment that has gone into these huge platforms, and focus on building an add-on ecosystem on top of it that focuses on protection of the interests of farmer?

2. Co-creation and data sharing: state of the art

This chapter discusses the need for public involvement and co-creation in the age of digital innovation, with an emphasis on the role of data. How is data collected and how do research results lead to impact in the city? And what does it take to scale up and sustain data-driven collaborative projects, such as the examples given in the middle range of Table 1?

The first section asks the question what we can expect from the 'governability' of the digital age. Digitalisation brings with it both advantages and disadvantages, which we need to be well aware of—otherwise our assumptions may be overly optimistic (or pessimistic). The conclusion drawn here is that with rising complexity, both of technology and of society, intricate models of collaboration are required to spawn off proper solutions. The second section, therefore, introduces the conduct of co-creation and explains how 'living labs' allow stakeholders to work together in constructive ways. Section three discusses how, in a society with high ambitions for participation, data is a pivotal asset, but we are only beginning to learn how data can fulfill these expectations. By talking about *data commons*, we might improve our understanding about what we need to share and in which ways. These insights are then made concrete in a proposed 'process flow' for collaborative data projects.

The empowered, digitalising society has its pitfalls

A lot of data goes around in the city. This brings opportunity: because of the low costs of replicating and disseminating information, digital technology can have a catalysing effect on the initiatives of CSOs, researchers and other (urban) stakeholders. As a result, many new innovations that tackle social problems are developed *online*, using apps, platforms and data systems. Areas where data and digital technology are used to meet societal challenges are diverse and range from healthcare and education, to democracy and the environment.[2]

But there is a flipside to this. As much as digitally-driven collaboration is *possible*, it has also become *necessary*. In our modern, borderless, digitalised society, intensive collaboration and coordination are increasingly a prerequisite to solving challenges. Problem owners now fiercely depend on real-time exchange of information and working with flocks of stakeholders in order to get things done. This is due both to the growing *interrelations* between current societal challenges (e.g. the connection between urbanisation, environment and healthy aging), and the *changing governance patterns* brought about by digitalisation (principally decentralisation and collabora-

---

2   See the DSI4EU thematic reports on digital social innovation in six thematic clusters: health and care, skills and learning, food, environment and climate change, digital democracy, migration and integration, and cities and urban development: https://digitalsocial. eu/blog/118/mapping-dsi-how-tech-is-tackling-society-s-biggest-challenges

tive governance). In sum, the big society we are creating thanks to increasingly powerful ICT is getting smarter, but also more unwieldy. This can be described as a 'rebound effect' of digitalisation.[3]

Along the same lines, much of the democratic potential resulting from digitalisation can be offset by the complications and pitfalls it produces. The increased availability of information and possibility to make oneself heard has given rise to a more involved, and more critical, citizenry. The means through which this is organised may suffer shortcomings, however. The processes and platforms by which citizens are involved may lack quality and guidance; they may also amplify existing imbalances and inequalities, or even produce new ones. Think of phenomena like fake participation, filter bubbles, people who are more complaining than contributing, or people being flooded by information (not even talking about its quality or veracity). In sum, this could amount to a relative *loss* of power instead of an increase. It means we are not creating a public sphere where voices are heard and people have voice, but produce fragmentation and frustration.

This may sound like a bitter warning, and one addressed to large companies and governments at that. But these issues are found in small, volunteer-driven projects just as much. Especially small projects that are aimed at bringing positive value should think how they can build a large and accessible impact. How do you expand beyond a

small circle of enthusiasts with pre-existing know-how? If we want to make the most of collaborative, data-driven and citizen-centric problem solving, appreciating people's life worlds is decisive. In order to not only be sensitive to people's contexts, but also genuinely involve them, co-creative methods offer a solution.

Co-creatively addressing urban challenges

As has been argued above, the urban challenges discussed in this report are generally too complex for a simple top-down approach. They may feature many interdependencies, stakeholders, or a dynamic context which is subject to sudden change. The strategy of co-creation, and the 'living lab' model in particular, is suitable to deal with these circumstances.

Co-creation is a term that denotes multi-stakeholder processes in which participants exchange views and ideas and look for 'the question(s) behind the question'. Practically, co-creation has people of different backgrounds coming together in work sessions, and collectively go through a series of steps. Waag's Co-Creation Navigator (Fig. 2) defines five zones or stages that contain the possible steps a co-creative project can go through. The co-creative process aims to yield outcomes that are valued by the different stakeholders present. These outcomes could be new ways of framing the problem, ideas of addressing it, and potentially designs of solutions that could work. At the minimum, co-creation will provide a better understanding of the problem you're trying to address. One bottleneck could consist of not knowing the motivations of the different stakeholders, particularly those that might not be able or willing to articulate their position; through co-creation, that bottleneck may be tackled. At the apex, co-creation will ensure that your project has buy-in from every stakeholder involved.

---

3   This is an expression often used in the context of analysing energy savings and other sustainability measures. "In conservation and energy economics, the *rebound effect* (or *takeback effect*) is the reduction in expected gains from new technologies that increase the efficiency of resource use, because of behavioral or other systemic responses. These responses usually tend to offset the beneficial effects of the new technology or other measures taken." From the Wikipedia entry on 'Rebound effect', accessed May 10, 2019.

Fig. 2. *The Co-Creation Navigator is a tool developed by Waag and partners to help with setting up and running co-creative processes. It contains toolkits and materials for warm-up games and other activities. Though co-creation isn't necessarily a linear process, there definitely is a sequential logic to it. The Navigator is set up as a journey through the co-creative landscape, using the metaphor of a subway map to zone out the different stages of the process.* http://ccn.waag.org.

Urban challenges often require stakeholders to commit for a longer duration: co-creation for the long haul. **Living labs** are a recognised means to achieve this. A living lab is a space for innovation where people share the conviction that they can or should develop their solutions in collaboration.[4] This means in concrete terms that the participants exchange experiences, test strategies against each other and share expertise. AMS Institute specialises in developing Living Labs through tools and resources that are to be found on their website.[5] In particular, researchers of AMS Institute have published *The Living Lab Way of Working*, consisting of eight steps that have co-informed the process flow proposed in this report (and presented on page 18).[6]

4   According to the European Network of Living Labs (ENoLL), living labs are "user-centered, open innovation ecosystems based on a systematic user co-creation approach in public–private–people partnerships, integrating research and innovation processes in real life communities and settings". See https://enoll.org/about-us/

5   https://www.ams-institute.org/how-we-work/living-labs/
6   https://www.ams-amsterdam.com/wordpress/wp-content/uploads/AMS-Living-Lab-Way-of-Work-print.pdf

The digital era requires more and deeper civic participation

Section 2.1 already spoke to it: the interplay between digitalisation and societal problem-solving. 'Digital technology' may sound like the type of solution that, once invented, is applicable everywhere, but often quite the contrary is true. The question who is shaping the technology has become increasingly relevant in a time where digital technologies influence and alter cities all over the world. While new 'smart' technologies can offer solutions to social challenges, citizens are rarely involved in such processes of innovation. Moreover there is a risk of coming up with 'techno-fixes'. The emphasis put on technology by governments and corporate actors can lead to stimulating high-capital tech investment at the cost of overlooking more subtle social processes and opportunities.

Digital social innovation aims to level the attention given to technology vis-à-vis the intricacies of society and social life. It acknowledges that innovation is complex, muddy, and brings forth positive and negative spillovers to different parts of society. This applies squarely to innovation in the digital age, which has technology branching deeper into people's everyday lives than earlier technological revolutions.[7] To ensure that technology addresses issues relevant to communities, the communities themselves should decide what technological solutions are best to use and how this should be done, for example through co-creation. By involving concrete communities in the design and implementation of technology,

citizens are empowered and negative indirect effects of technologies can be largely avoided. Of course, community involvement in technological development is not only relevant for city dwellers: the organisation FarmHack, for example, brings together hacker, planner and farmer alike to make sure the novel world of 'agritech' and 'smart agriculture' unfolds in the benefit of society with applications that match the needs of the farmer (see the interview with Anne Bruinsma, page 11).

In order to succeed, digital social innovation requires people to be able to genuinely collaborate through the means of digital connectivity and online platforms. With the major social media platforms seem to have degenerated into advertising and data harvesting machines and value-driven alternatives struggling to get a foothold, there is a challenge ahead of us. How do we properly develop open, cooperative platforms that spur collaborative action and the sharing of knowledge, experience and solutions? How to encourage long term engagement of the user community? These questions are not straightforward, and depend—we argue—on precise identification of communities' needs in well curated processes.

Data as the nub of innovation and collaboration

Cities are highly data-rich environments. Countless sensors, smart devices, open access platforms and data collections provide a potential resource for use and re-use. But most of that data is locked in proprietary systems and constrained for limited purposes. It can be quite impossible to share the data based on concerns of security, privacy, commercial considerations, or because of practical concerns (in the case of huge, or very specific, datasets). Official censuses and statistical records (such as those published on data.amsterdam.nl) entail a

---

7   J. Rifkin, *The Zero Marginal Cost Society: The Internet of Things, the Collaborative Commons, and the Eclipse of Capitalism*, Palgrave Macmillan 2017; S. Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, Profile Books Ltd 2019.

more directly useful and relevant source of data. Similarly, the knowledge produced during the conduct of academic and civic research projects is of great potential value. Fellow researchers, students, citizens, governmental parties, and companies operating in the region may derive insights from the methods employed, knowledge derived, or through direct access to the underlying data. However, project partners often face challenges in terms of secure data storage throughout the research activities, and sustainable access to outputs after a project is completed. Therefore the main question is how to bring together the different skills and assets that are needed to sustain a collaborative, data-driven project; and to do so in a way that maximises the opportunity to work together on urban challenges. Put in other words: how do we organise the 'knowledge question' as a public good?

The perspective of **data commons** provides a clear direction. Data commons is a citizen-centric approach to data governance.

Data commons seek to generate sustainable value by sharing data in common, whilst giving people control over their personal data (as required per the General Data Protection Regulation (GDPR)). While driven by the ambition of opening up code and data, a data commons approach does *not* plead for opening and sharing *all* of our data. It rather recognises (1) that most data is sensitive and (2) data can be put both to good uses and bad ones. Because of these concerns, putting a *commons* lens on managing data is merited (see next page). Their value-driven incentive model makes that self-ownership is more likely to be supported in a commons rather than in a privatised model.

In the next section, we will discuss how the insights about data commons can be put to practice. If we want to have an inclusive society, with high ambitions for participative and data-driven governance, how should we organise our processes?
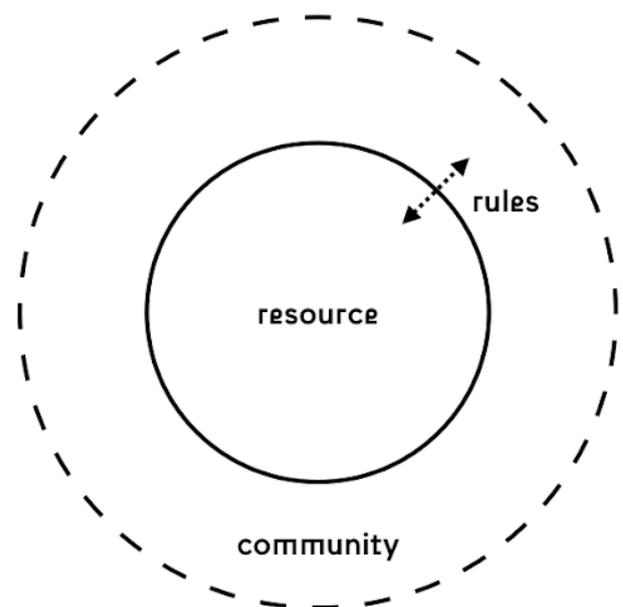
What are commons?

Commons are shared *resources* managed by empowered *communities* on the basis of mutually agreed and enforced *rules* (see image). Think of a food cooperative which brings together consumers and farmers and helps them build a food system based on fair pricing, healthy farming and social cohesion.

Commons need to be protected, rather than be left open for anyone to exploit. Proper commons are organised through legal entities such as cooperatives, foundations and associations. There are commons initiatives for the management of natural resources, social services, cultural heritage, for keeping streets liveable, for spreading knowledge, code and design, and so on.

Some key characteristics of commons:

- Commons strive to achieve core values 'by design' (sustainability, inclusiveness, privacy, accountability), as opposed to compensating for negative side-effects afterwards. A positive energy housing block is 'sustainable by design' for example because it generates a surplus of energy and does not rely on other, external sources to supply energy or compensate emissions.

- Commons thereby assume a direct involvement of peers (i.e. citizens, practitioners, beneficiaries) in, and even co-ownership of, the resources and solutions.

- Communication is crucial to sustain commons. Determining values and deciding on what is needed is not based on technical assumptions, but through conversation and mutual trust. These must be insured or supported by means of rules and agreements.

- Community initiatives often look for new forms of cooperation and (financial) support. They are not easy to catch in existing categories and that sometimes makes it difficult to move forward.
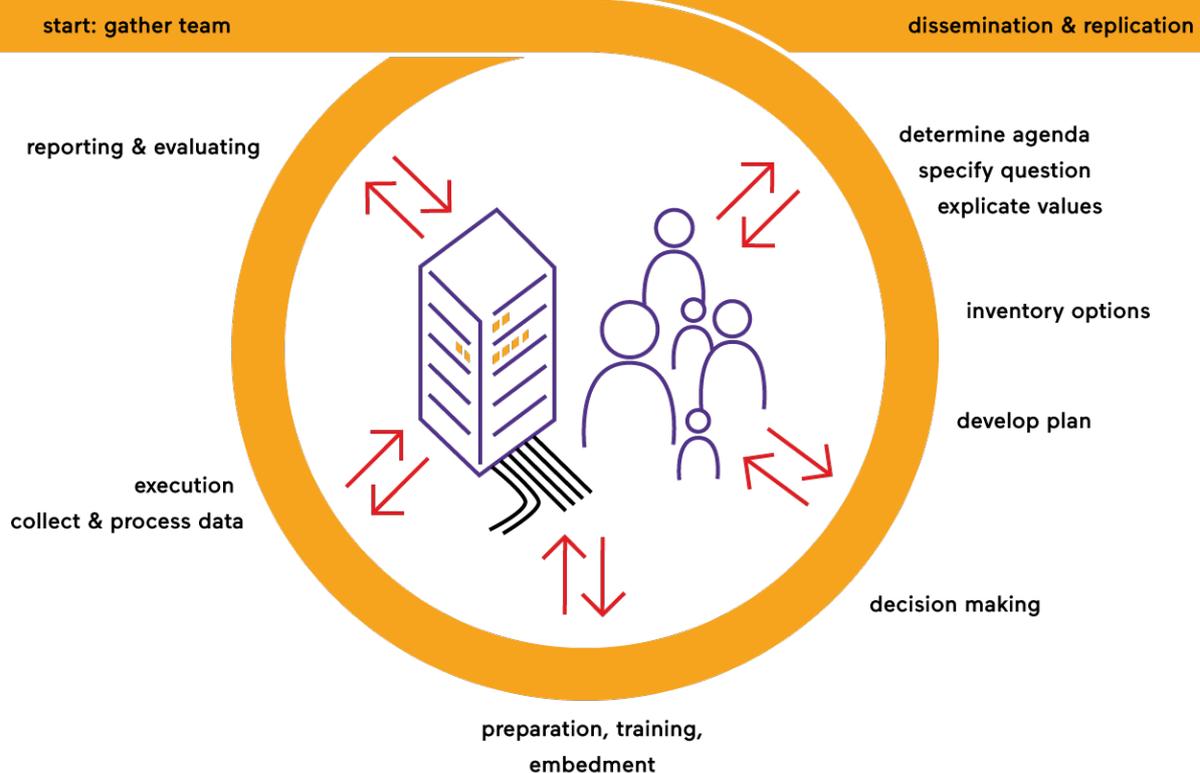
The process flow

The smart society derives its smartness not from a large number of devices and sensors that generate data, but from a learning and development process that is open to a large number of parties. More important than building huge data stores is therefore the creation of a complete data process in a trusted environment. The collection of data for knowledge development and management must be regarded as a process, in which it is constantly discussed why data is collected and under what conditions. This keeps the main question in the foreground and makes the initiative more accessible to new parties, insights and solutions.[8] The process flow for data commons lays out the steps required to continually adjust to the crucial ingredients of the data commons. Think of defining the community, the why-question, the collection of data, the decision-making and the way managing and valorising data. This process flow has been drawn up by analogy with policy, research and project processes.

The composition of a well-defined group of parties is of particular importance for collaborative data projects. This can involve a group of public and private partners, but also a community of citizens. This requires a trusted environment. This is more than a place where all participants feel at ease: It is the formation of a group of participants who feel mutually familiar with each other and make agreements about the nature and form of the collaboration. Within this environment, people share the conviction that they can or should develop further and are prepared to work together on this.

The process should be approached as an integral way of working and be provided with adjustment options. That way the group can keep innovating and learn from every process step.

start: gather team — dissemination & replication

reporting & evaluating

determine agenda
specify question
explicate values

inventory options

develop plan

execution
collect & process data

decision making

preparation, training, embedment

---

3. Lessons and recommendations

This chapter discusses the main consequences of the process orientation on collaborative data projects proposed in chapter 2. In the first section, we discuss four assumptions that need to be reconsidered. Next, we outline how 'system players'—both intermediate organisations and public institutions—can take up leadership in advancing this kind of data collaborativism.

Overcoming problematic assumptions

The process view on collaborative data projects helps to keep the goal central and to avoid certain persistent assumptions.

The **first** typical assumption is that data collections only have to be set up and will then work smoothly, meeting many users' needs at negligible marginal cost. The requirements for cooperative data projects are specific to each case and situation. Not the data technology and the existence of a server are decisive ingredients, but the maintenance of the data, the bundling of knowledge (bringing together additional specialisations), and mediation between those roles is what it's all about.

A **second** assumption is that 'data sharing' is by itself a tempting premise for companies and other parties. However, the possession of knowledge and information is inseparably part of the business model. Once the feasibility of a data-driven solution has been determined, such as logistics optimisation, it will take a while before profitability is achieved. Therefore, a 'closed' approach is initially safer than an 'open' approach. When are the costs and risks of disclosure outweighed by the benefits of shared

knowledge? Without a clear question or use case and trust in rewarding outcomes, it is difficult to have the confidence to make the leap to data sharing.

A **third** assumption is that databases are neutral. Data are—contrary to what many people think—not objective: every data point is an interpretation of reality. Every data collection entails a choice about what is 'part of it' and what is not. A single data point is not an isolated truth, but it is a snapshot or a measurement that acquires meaning in its context. The design of a database should be a translation of the (social) context in which the data is collected, but that is not always the case. Databases are often set up with a specific party in mind and optimised for their interests. Prejudices may creep into the system that advantage certain parties better than others. Much of the art of getting knowledge from data lies with the 'metadata': information that provides information about other data.

A **fourth** assumption is in line with this: the idea that 'open data' is really open and accessible to everyone (and similarly, that 'shared data' is really accessible to every participant). Organising and interpreting data takes time and money and requires certain specialisms. In addition to being able to handle a data set, a user must also have knowledge of the context and must then be equipped to do something with that knowledge. Having access to data does not immediately lead to a better knowledge position. A certain knowledge position is needed *at the start* to have actual (actionable) access to that data.

Now that we have a better understanding of the pitfalls of collaborative data projects, and we have learned about the 'data com-

mons' approach and the necessity to work from a process model, how can we put these insights in place? And *who* are the key players that can help make this work? The next section concludes this report by synthesising the lessons into recommendations.

The role of intermediaries

This section presents our recommendations, geared especially towards organisations that want to support civil society with their shared data management. As analysed earlier in the DSI4EU project, the digital social innovation community in Europe consists of a significant number of network 'nodes' that act as connectors, facilitators, knowledge repositories and physical spaces for collaboration, among other roles.[9] These are 'third sector' organisations, i.e. non-governmental and non-commercial organisations with societal objectives, with a particular 'connective' skill. These organisations may take on a range of roles, but share the characteristic of being network shapers. They are able to maintain active relationships with, and establish connections between, DSI practitioners, CSOs and institutional players.

As explained in this report, collaborative data projects depend on active relationships, knowledge transfer, and a shared practice in order to thrive. Intermediary organisations such as those mentioned above facilitate these conditions. The ITSSOIN project ('Impact of the Third Sector as SOcial INnovation', EU FP7 no 613177, 2014-17) studied the nexus between the third sector and social innovation, helping to develop understanding of why and how the sector is central to functioning, cohesive and viable societies. The project's main hypothesis was that the third sector, specifically through stimulating civic involvement, is best placed to produce social innovation, outperforming the commercial sector and the state in this regard.

The ITSSOIN project gathered valuable insights about the particular qualities and functions these organisations typically demonstrate. Citing from their project reports: "Favourable for taking this [intermediary] role were, among others their proximity to target groups, their long or specialised expertise in working with those groups, a sense of devotion to a cause, their trustworthiness, openness and high degree of social capital, in particular embeddedness. Such non-profit 'hubs' frequently invited partners to contribute and thereby acted as a bridge into a collaborative constellation. The 'connective' action of third sector organisations, which often seemed to even go beyond their ability to perform 'collective' action, deserves closer attention, in particular as regards their specific function in social innovation processes".[10] What emerges from these findings is that intermediaries can fulfill their connective role because they are relatively 'close' to practitioners both in a physical sense and by their nature. This is a unique feature which is hard to substitute for by other (more formal) institutions. However, as the next section will show, these 'system actors' also have a concrete and important role to play.

The role of system actors

Governments and other established institutions (here also referred to as 'system actors') are often further removed from the (life) worlds of those involved, and also maintain a more formal relationship with

---

9   Schouten, S., Gerritsen, M., and Jansen-Dings, I. (2018). Analysis of DSI super node methods – A deliverable of the project DSI4EU. Amsterdam: Waag

10   Cited from p. 55/56 of D8.2, "The Impact of the Third Sector on Social Innovation" of ITSSOIN project.

them. Professionals in these institutions may be less able to 'connect' with the target group, although government philosophy has recently shifted in favour of building direct relationships. According to the Dutch School of Public Administration (NSOB), governments have shifted the emphasis away from administrative and managerial approaches towards modes of 'networked' and 'participative' governance.[11] Through these roles, but also through their more traditional governance roles, system actors may importantly contribute to the success of collaborative data practices and give rise to 'data commons'.

Importantly, governments can set public standards, such as technical and procedural norms for data sharing. This entails a passive governance measure. Governments (and other public institutions) can also nurture a rights-based approach, in which citizens and CSOs are *actively* facilitated to take part in collaborative data projects. As we saw in the first section of this chapter, one of the problematic assumptions in open and shared data projects is that access is a passive feature: opening a data set is enough. But that is far from true in most cases. Certainly citizens and civil society organisations need forms of active empowerment in order to be able to take part in data projects.

Of course, the level and extent of involvement of certain (non-expert) stakeholders will depend on the specific ambitions of projects, which may be limited and only require a basic level participation. Awareness and engagement projects, for example, target the wide audience, which have generally a limited skill and motivation, but do have potential for activation and learning. But citizens and CSOs may also be involved with the aim of amplifying research efforts, in which case highly skilled individuals and groups of individuals are targeted. Another objective may be to stimulate the (political) empowerment of communities, such as in participative environmental monitoring or a community-led food transition project. In all of these cases, system actors that want to support collaborative data projects, have scope for both passive and active methods of facilitation and standard setting.

The take-away for city administrators is thus to ensure that data-driven initiatives may arise at different levels, but that they always have an 'open process orientation' (such as proposed in the process flow in the previous chapter). That way, cities can absorb new ideas and people, be open to different levels of specialisation and tap into many sources of creativity and problem-solving capacity.

---

11   Teisman, M. et al., "Effectief Sturen met Multi-level Governance", NSOB 2018, p. 31.

Our city data strategy, in sum

Digitally innovative cities have great potential to address wicked challenges, ranging from unemployment to traffic congestion and water quality. But as this report has sought to show, the tree of digitalisation bears fruits that may be ripe and unripe, hang low and high, be healthy and unhealthy.

Many cities have high hopes for data-driven solutions for their problems, and expect that these will also allow solutions that are more inclusive and participative. In order to fulfill these promises, a shift of focus from 'building data hubs' to organising a more inclusive process is necessary. The four steps below summarise the way this reports suggests this could be done.

1. **Initiate:** Start with a well-defined team of organisations that share an ambition. Agree to work together as peers to tackle a problem, create a joint (data) market, or further (scientific) knowledge (or all of these simultaneously).

2. **Scope:** Identify the level of 'user' knowledge and specialisation (taking for example Table 1 as a starting point). From that, the group can determine how big and involved the project will be, and what degree of 'institutionalisation' (collaboration with, or dependence from, system actors) may fit.

3. **The process:** Make sure to follow a thorough process routine, and to open that process (at least at some points along the routine) to everyone who wants to participate. Note that openness to new entrants does not mean everybody can just walk in and out: careful curation of the process is warranted.

4. **Set a standard:** If the process works well, share your way of working. When projects share the same process standards and comply to a minimum of openness, they are well suited to be disclosed and shared on generic collaborative data platforms—platforms that are not designed to accumulate data, but to allow for active participation and knowledge sharing.

About AMS Institute

Amsterdam Institute for Advanced Metropolitan Solutions (AMS Institute) is a public-private institute founded in 2014 by Wageningen University & Research (WUR) and Delft University of Technology (TU Delft), together with Massachusetts Institute of Technology (MIT). AMS Institute is an internationally leading institute where talent is educated and engineers, designers, and both natural and social scientists jointly develop and valorize integrated metropolitan solutions.

AMS Institute's mission is to develop a deep understanding of the city – sense the city – to design solutions for its challenges, and integrate these into the city of Amsterdam. This is done through three main activities: education, research, and urban data.

Education: The Master program Metropolitan Analysis, Design and Engineering (MSc MADE) aims to educate urban engineers that are equipped to deal with the metropolitan challenges of our cities. In its interdisciplinary courses, students attain skills and knowledge by testing their ideas in projects and living labs in Amsterdam.

Research: The research and valorisation activities revolve around six urban challenges: Smart Urban Mobility, Urban Energy, Climate Resilient Cities, Metropolitan Food Systems, Urban Data & Intelligence, and Circularity in Urban Regions. Each challenge is addressed with collaborative research to develop a deep understanding of the city. Integration of these challenges provides opportunities for creation of an innovative, sustainable and just city.

Urban Data: Data are at the core of understanding the city, and help with urban planning and decision making. AMS Institute designs, develops and studies methods and tools for urban data processing, visualizations, and applications.

About Waag

'*Make technology & society more open, fair and inclusive.*'

For over twenty years, Waag has operated at the intersection of science, technology and art. It focuses on emergent technologies as instruments of social change, guided by the values of fairness, openness and inclusivity. Waag's dedicated team of sixty thinkers and makers empowers people to become active citizens through technology.

With *public research*, Waag proposes a third research domain that encompasses appropriate scales and actors needed to further social innovation. Independent from the market and the state, it understands innovation as a process of interrelated change in the technological and socio-ecological realms.

Public Research starts from the premise that society is a research community. This premise has a number of important consequences that make it distinct from, yet open to scientific and industry research practices. Public Research is fundamentally interdisciplinary, being performed by heterogenous and phenomenon-specific research communities. For these communities, the goal is not to articulate the matters of fact of scientists, nor the matters of interest of industry, but rather the matters of concern of citizens. The multiplicity of approaches and findings that such research produces, implies a fundamental complementarity of different research methods, communities, and their outcomes. Deciding what matters of concern, what methods and outcomes are relevant in studying a certain phenomenon is for Public Research not a question of objectivity, but rather one of ethics. The question is no longer what is true, but who and by what procedures is allowed to produce knowledge in a democratic society.

Colophon