

Self Data cities, the playbook

“WHAT IF CITIES TOOK A CENTRAL ROLE IN RETURNING CITIZENS’ PERSONAL DATA TO THEM?”

THE “MESINFOS - SELF DATA CITIES” TEAM AT FING:

Manon Molins, Chloé Friedlander, Guillaume Jacquart, Fanny Maurel // Sarah Medjek for MyData France.

TRANSLATION: Jianne Whelton

CREATIVE COMMONS

This document is available under a Creative Commons Attribution 4.0 License (France): <https://creativecommons.org/licenses/by/4.0/deed.fr>.



You are free to share – copy and redistribute the material in any medium or format – and adapt – remix, transform, and build upon the material for any purpose, even commercially – under the following terms: attribution – you must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests that Fing endorses you or your use. You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

FING IS SUPPORTED BY



ACKNOWLEDGEMENTS:

Special thanks to Virginie Steiner (La Rochelle), Sylvie Turck (Nantes Métropole), and Maria-Inés Léal (Grand Lyon), as well as to Guillaume Chanson, Cécile Christodoulou, Aurialie Jublin, and Mathilde Simon.

What if cities took a central role in returning citizens' personal data to them, so their citizens can use personal data to make their lives easier, get to know each other better, contribute to territorial decision making or participate in public interest projects?

For the past year, Fing has been working with three major French cities – Nantes Métropole, (the energy transition), La Rochelle (sustainable mobility), and Greater Lyon (social welfare) – to enable them to implement a Self Data experiment of their own as early as 2020. This playbook is the fruit of our work together. Its central objective is to put Fing methods into the hands of other metropolitan areas looking to dive into Self Data, so they can do so feeling suitably equipped. In this kit you will find:

> an introduction to the notion of Self Data – sharing personal data with the individuals they pertain to – and the major issues that Self Data implies for cities;

> an analysis of the relevant governance models when considering how to share personal data with individuals;

> a survey of cities efforts to share data, including some examples to draw inspiration from (and some to avoid);

> illustrated methodologies you can use to implement a Self Data initiative in your region (plus examples drawn from our work with Nantes Métropole, La Rochelle and Greater Lyon): identify the relevant personal data, imagine use cases that utilize these data, look towards governance models to frame your use cases, and select your experimental scenario(s);

> some advice born of our experience working with Self Data since 2011, which we would have loved to have ourselves when we first started experimenting with Self Data at Fing.

Happy experimenting with Self Data in your region!

Sommaire

01

02

03

04

Annexes

"MESINFOS: SELF DATA CITIES" – WHAT IF CITIES TOOK A CENTRAL ROLE IN RETURNING CITIZENS' PERSONAL DATA TO THEM?

- P - 02 >>** Self Data: the concept
- P - 07 >>** Can Self Data be a lever for metropolitan innovation?
- P - 09 >>** Sharing personal data: governance models

"CITIES AND PERSONAL DATA" – OVERVIEW OF GLOBAL INITIATIVES

- P - 20 >>** Cities and personal data, what ways ahead for citizens?
- P - 26 >>** A closer look at cities' personal data initiatives

"SELF DATA DIY" – CREATE A LOCALIZED, RELEVANT SELF DATA INITIATIVE FOR YOUR REGION: THE EXAMPLES OF THE CITY OF NANTES, LA ROCHELLE, AND GREATER LYON.

- P - 40 >>** Understanding the stakes associated with Self Data and choosing an appropriate project
- P - 42 >>** Defining the data parameters
- P - 43 >>** Potential use cases and relevant models of governance
- P - 66 >>** In practice: outline an experimental scenario for your region

"SHALL WE BEGIN?" – PROTIPS FROM EXPERT SELF DATA ACTORS

- P - 74 >>** A Self Data experiment: 10 questions to ask
- P - 76 >>** A Self Data experiment: 10 levers to activate
- P - 78 >>** A Self Data experiment: 7 pitfalls to avoid
- P - 80 >>** A Self Data experiment: 10 cardinal rules to respect

APPENDICES – READY-MADE METHODOLOGIES

- P - 84 >>** 1) The seminar
- P - 88 >>** 2) The datablitz
- P - 92 >>** 3) The use cases
- P - 96 >>** 4) The governance models
- P - 100 >>** 5) The experimental scenarios

01

19

39

71

82

An aerial photograph of a city, showing roads, buildings, and green spaces, is overlaid with a semi-transparent blue filter. A large, white, outlined number '01' is positioned in the upper right quadrant of the image. To the right of the number, there is a block of white text.

01

“MESINFOS: SELF DATA CITIES” – WHAT IF CITIES TOOK A CENTRAL ROLE IN RETURNING CITIZENS’ PERSONAL DATA TO THEM?

01

“MESINFOS: SELF DATA CITIES” – WHAT IF CITIES TOOK A CENTRAL ROLE IN RETURNING CITIZENS’ PERSONAL DATA TO THEM?

SELF DATA: THE CONCEPT

For more than seven years, within its wider MesInfos research trajectory, Fing has been working on the subject of **returning personal data to the individuals they concern** – a practice and concept we have dubbed “Self Data.” After crossing what amounts to a major threshold, it seems like the right time to pass the torch on to you.

Our goal has never been to make Self Data a proprietary concept that only we hold the secrets to. On the contrary, we have always wanted to share our work and enable personal data initiatives to multiply. Seven years later, it’s obvious that although the concept remains unique, Self Data is no longer the unknown concept of yesterday. This kit aims to democratize it further via the efforts of key players that, we believe, are uniquely capable of propelling Self Data dynamics: cities.

Self-data – which we define as “the collection, use and sharing of personal data by and for individuals, under their complete control and designed to fulfil their own needs and aspirations” – implies that individuals should be the masters of their data, and that they should be able to do things with it that serve their personal needs. For this to be possible, individuals must do three things:

1) Obtain a copy of their personal data #API #dataportability #GDPR

The keystone of Self Data is enabling individuals to obtain a copy of their personal data, which at present is stored in various organizations’ information systems (energy providers, telephone operators, social networks, etc.). These “data holders” must, therefore, build the necessary data channels. This right to data portability – making personal data available to “data subjects” – is guaranteed by Article 20 of the General Data Protection Regulation (GDPR), which came into force in May 2018. Data portability, an essential component of the Self Data concept, necessarily applies to any type of personal data. Returning individuals’ personal data to them will make it possible for them to get a 360° overview of not only the tasks they perform online, but also all the data they generate doing their shopping, speaking on the telephone, sending a text message, taking the car, turning on the lights, etc. It is in the interest of every single data holding organization to make Self Data a reality!

2) Personally store and manage their data securely. #PIMS #PDS #PODs

The second Self Data imperative: make it possible for individuals to store and manage their personal data securely. This may be by offering them a personal digital home: an online, personal space

they fully control, where they can gather together (aggregate) their data. This would give them their own information system similar to organizations’ CRM systems, and increasingly powerful tools they use to manage and analyze the data they hold. Although there are not many, such tools are emerging steadily. They may take the form of a PIMS (Personal Information Management System), whose market contenders include Cozy, Digime, or Inrupt by Tim Berners-Lee (in connection with his SOLID project); they might resemble a digital “safe” that can integrate data sharing logic (like France’s DigiPoste); or perhaps the space is offered by trusted actors, Fair & Smart and Onecub for example, who position themselves as guarantors of the transfer of personal data from one data controller to another.

3) Reuse their data to enhance their everyday lives, contribute to public and collective life, etc. #thirdpartyproviders

Once the data has been retrieved and aggregated securely in the personal space, it’s time to use it – or rather, reuse it! Only through use does the full value of the data reveal itself. Early adopters and (social) activists are already testing organizations’ ability to respond to a portability rights request and generate data in a machine-readable format, while others with

technical skills are developing personal tools they can use to visualize and correlate their data and establish correlations. But for the majority of individuals, third-party services are the means with which they will make sense of Self Data. Therefore, third-party reuser services that enable individuals to derive value from their data are needed. After having explicitly asked for individuals’ consent, third-parties would process individuals’ private data and offer them relevant services – without necessarily having access to the data themselves.

So what *can* people derive from their data? Why would individuals even want to see Self Data emerge, and how might they benefit from it individually or collectively? During our MesInfos years, we have explored a wide variety of personal data use cases, which we have grouped into seven categories: making life easier; better controlling who has access to what information; getting to know each other better; living in closer alignment with personal values (greener, more ethical consumption); contributing to shared knowledge creation; making well-informed decisions; or simply changing things up. People could, for example:

» get smart(er) bank statements that make their lives easier. Each line of their bank statements would be placed into a meaningful context if triangulated with data

from other sources (invoices, Social Security), which might help people avoid unwelcome financial surprises;

» cross food consumption data with a carbon benchmarking system, such as Ademe’s Base Carbone®, to assess their diet’s carbon footprint, or with their health data (allergies, specific diets, etc.) and Open Food Facts to create shopping lists more closely tailored to their needs;

» identify the vendor contract offers that are most suited to their individual needs according to the data in their profile, plus easily switch telcos, power suppliers, insurers, etc.

The ideas and concepts emerging are numerous. These efforts will finally grant individuals some form of mastery over their data, which could reestablish trust and redress the balance of power between organizations and individuals. After all, our notion of Self Data from the perspective of organizations is, “If we can use your data, you can too... however you please.”

Today, very few people can do those three things. Nevertheless, **there have been many promising developments in the realm of Self Data in France and elsewhere in the world.**

» Fing – which launched the Self Data re-

search track in France – has been carrying the torch since 2011, and its [MesInfos project](#) is now in its sixth year, thanks to the support of its partners and allies. We began by exploring Self Data and looking at what others were doing elsewhere (the Blue and Green Buttons in the United States, MiData in England, the VRM community, etc.), before testing the return of personal data with 300 participants in 2013 – a worldwide first. We worked on the technical, economic and legal challenges related to Self Data, and then specifically on the sharing of energy and health data. The issue of health-related self-data we then spun off into the "My Data, My Health" project. Naturally, we have spent time looking closely at the right to data portability. We have worked closely with members of our Dataaccess working group, and across an expansive Pilot study that took place between 2016 and 2018, where we collaborated with over 2000 testers and multiple data holders to demonstrate the scaling potential of Self Data.

» Internationally, MyData has evolved into a worldwide community of collaborators, all working to recalibrate the benefits generated by the personal data economy more towards individuals. In October of 2018, the network – launched in 2015 by a small group of interested actors including Open Knowledge Finland and Fing – was reborn as [MyData Global](#),

an officially-recognized Finnish non-profit comprising more than four hundred members. Fueled by the wider goal of making individuals the masters of their personal data, the MyData movement became a fully-fledged non-profit organization based in Helsinki, with twenty local hubs around the world. Its annual conference – whose [4th edition](#) takes place in 2019 – acts as its principal catalyst.

» The key to Self Data lies in the hands of the data holders. They establish the proper channels for data transmission and facilitate data reuse when they observe the correct documentation practices and work to generate synthetic datasets. Without these, Self Data would be but a concept, kept alive thanks to vaguely permissible data scraping techniques. Today, very few actors are jockeying for position. Some of digital's heavy hitters – the ones with longstanding data transmission channels like Facebook and Google – are launching initiatives and forming coalitions of actors to address the right to portability ([Data Transfer Project](#)), or seeking to position themselves as experts on the subject, ready to give advice to other sectors (see Facebook's report from September 2019, entitled "[Data Portability and Privacy](#)"). There are, however, some data holders who are truly invested, and who are leading by example.

» Energy providers Enedis and GRDF have set up processes and projects (Enedis Data Connect, GRDF Adict) to allow individuals with a connected meter (Linky, Gazpar) to recover their data and reuse them – to identify better contracts, better manage their consumption, etc. – via third-party services. These companies presently form the vanguard among data holders. Their position is rather unique: because energy code regulations required them to contemplate data sharing long before the right to portability came into effect, and given their role as energy distributors – not service providers – they were more inclined to share data directly with individuals in ways that enabled a third-party service market to emerge.

» Other organizations and sectors are making the right to portability a concrete reality for their users. An initiative started by French online classified advertising giant LeBonCoin is worth mentioning, for its rarity: "data takeout is a feature that allows you to download and view your personal data autonomously". According to [Damien Dégremont](#), Data Protection Officer at LeBonCoin, "Since May 25, when the feature became operational, we have been receiving several thousand requests for personal data per week, versus the handful we were receiving per year before it." Curiosity is spreading . . ! Enabling individuals to reuse their personal data via third-party services is

a way to create value that establishes a virtuous cycle.

» Noteworthy also are the advances in data portability taking place across entire sectors, France's health institutions have long been calling for a "Blue Button à la française" that would facilitate the sharing of health data between organizations and individual patients, which would, of course, be linked to France's [Shared Medical Record \(DMP\)](#) system. With the [Digital Health Space \(ENS\)](#), voted into existence by the National Assembly in 2019, individuals might potentially have a platform from which to manage their consent to share data, their treatment and prescription records, and reuse these and other manually entered data thanks to trusted services and third-parties. The telecom sector has launched several initiatives related to effective portability procedures in conjunction with the [Data Portability Cooperation project](#). On the banking side, in England – where the world's first initiative for personal data portability, the MiData program, was born – the [Open Bank Project](#) lets banks share customers' data with third parties, under the control and for the benefit of those customers. The program is similar to the European "DSP2" directive. After a [rocky start](#), the Open Bank now seeks to demon-



strate the value of reusing what is shared by linking up with innovative companies. In France, a collective has published a white paper on portability, which they have presented to the National Assembly, and seeks to create a cross-sector consortium on the subject (a new form of governance).

» Other good news on the Self Data front: experiments – beyond those we conduct at Fing – are emerging everywhere. In Brittany, the Académie de Rennes is conducting a project [helping national education students and teachers to gain mastery over their data](#) throughout their learning and career trajectories. According to Olivier Adam, "a lot of personal data is scattered in the digital education ecosystem,

stored on different devices, recorded across several applications, and sometimes misplaced at Google". The academy is working on an experiment that will provide participants with personal clouds (Cozy Cloud) and a copy of their personal education data. More on this experiment in the next chapter.

» In line with the MesInfos pilot, insurer Maif now [provides](#) its members with a personal cloud, where it shares the details of their home insurance policies with them. Likewise, in the context of Fing's "My Health Data" health care and disease prevention efforts, the Vyv Group is concepting Self Data services that use personal health data and a solution implemented by Fair&Smart that enables data reversibility – meaning that individuals can withdraw their consent and have all their data is erased immediately. In terms of the services, an individual's blood test results, for example, can be used to augment their disease prevention file, be enriched after cross-referencing with air

quality data or other benchmarks, serve as a means of monitoring their the state of their health and overall, help them to get to know themselves better.

» There are also large-scale experiments happening abroad. Iceland's Digime and the Icelandic Ministry are setting up a [health data sharing service](#) for individuals, the Finnish Transportation Department is spearheading a [MyData pilot study](#) looking at facilitating mobility data sharing, and the Finnish National Education Agency is considering avenues toward sharing education data.

» Growth in the PIMS market is also very promising. In 2011, when MesInfos launched, not one was truly operational. Today, dozens of entrepreneurs are jockeying for position: Digime, Cozy, Matchupbox, Fair&Smart, Onecub, etc. Tim Berners-Lee – the granddaddy of the World Wide Web – has launched a personal data control hub/platform called Solid. Even Microsoft is jumping on the bandwagon, with its own PIMS, known as [Bali](#). Back in France, Facebook's ["startup garage"](#) is supporting the emergence of services at various points on the data control spectrum – and so are many public administrations! In October 2016, the European Data Protection Supervisor (EDPS) estimated that PIM systems represent a "para-

digm shift" that "enable individuals to regain control over their data", and [has explicitly called](#) for European Union support. Tenders for H2020 project proposals are appearing (eg: [Supporting the emergence of data markets and the data economy](#)) – most of which are "Self Data" oriented. Finland, head of the Council of the European Union at the time of writing, has embedded the "need to create a competitive data economy, where individuals are in the driving seat, promoting access, interoperability and the use of data while respecting the rights and privacy of individuals" in its official [program](#). And in England, the ["Smart Data Review: Putting consumers in control of their data and enabling innovation"](#) led by Her Majesty's Government [is seeking](#) to better understand the role they can play in data-driven technology and services development. Finland's State Development company Vake launched a Self Data-themed ["Industry Hack"](#) open innovation studio, and lauded the national Post Office for its newly-designed PIMS. Its next steps? Launching a series of MyData pilot studies that will enable thousands of Finns to regain control of their data.

What all these initiatives have in common is that they are tabling new accountability frameworks – or at least thinking about them. Long gone are the days of minor examples of isolated services and platforms. All these actors talk to each other,

exchange best practices, sometimes compete and sometimes cooperate.

CAN SELF DATA BE A LEVER FOR METROPOLITAN INNOVATION?

Our work on [MesInfos](#) over the last 7 years has made France a central actor in an incredibly powerful wave of digital culture development, and one of the pioneers in the field of personal data sharing. All the good news in the preceding paragraphs makes us happy (or maybe causes us to worry), but in practical terms, how can we make uptake of the concept more widespread? Between 2016-2018, we carried out a major [pilot project](#) that brought together data-holding organizations, a platform, the citizens and government of one region in France and an entire innovation ecosystem to concretely explore the hidden potential of Self Data. The data holder partner organizations enabled more than 2000 individuals to regain control of the personal data they were holding, so that the individuals could reuse it. . . however they pleased.

Even though, in our role coordinating the MesIn-

fos pilot during all these years, we have proven we are capable of implementing a Self Data initiative on a grand scale, we never intended to carry the Self Data torch indefinitely. **So who will take it? Who is capable of managing the scope and complexity of this kind of project, one that brings together private (sometimes competing) organizations, public entities, researchers, laymen, plus an entire innovation ecosystem? We have become increasingly confident in our feeling over these last few months of the pilot that cities have a special role to play in the future of Self Data.**

At a time when regional authorities are facing complex challenges on an unprecedented scale (environmental, economic, social, developmental), the "smart" revolution is emerging apace: there are more and more intelligent systems and service that optimize flows, reduce consumption, mitigate risk. . . Is there perhaps another way, one that is less integrated and more open, one that offers inhabitants and regional actors some space to explore? What if an entire metropolitan area adopted the Self Data concept, and then managed the flow of information, tools and methods to and from actors big and small, in order that they all contribute to making the project work, as a matter of public interest? Personal data are often produced in a metropolitan context; crossing them with local Open Data would create a powerful

source of value creation not only at the individual level, but also at the wider (public and private) metropolitan level. Self Data could be at the service of the region's energy transition, could generate answers to mobility questions, be the springboard for new (public and private) services for inhabitants. . . This is what a data commons might look like, where data is shared equally between all parties concerned.

Just as massive metropolitan and regional actors do, individuals would be able to produce, capture, collect and retrieve personal and non-personal data. They could choose to share them, use them for themselves or even contribute them to studies and knowledge production for the region. In short, based on their personal data, individuals would have the agency to participate in the emergence of new services – independently or collaboratively – play a part in planning their metropolitan environments and the weaving of cities' urban fabrics, contribute to political decisions, advance collective knowledge creation...

Regional actors also stand to gain from returning personal data to Individuals. People are becoming increasingly ill at ease about the ways public and private organizations are processing their personal data, while legislative restrictions are getting tighter and tighter. Mere compliance with the law will not be enough to

avert the ongoing crisis of confidence. There is an alternative, however. These actors could lay the foundations for a new relationship with users, customers and inhabitants. At the same time, by building new frameworks for data sharing under individuals' control, they could be part of the data revolution, and garner for themselves a solid position in the emerging data services market (platforms, applications, etc.). Self Data practices can transform individuals' apprehensions and enrich the traditional activities of regional actors (service creation, planning, knowledge production, etc.).

In urban contexts where services based on personal data are legion, and yet cities and citizens obtain little in the way of value from them, a Self Data approach might very well turn the tables – not only because it seeks to protect people's privacy, although that is crucial, but largely because it is sustained by individual and collective empowerment through data use.

After more than 10 years of work at Fing on cities & digital – from our 2008 "Cities 2.0" program, to the more recent "Audacities: Governing a real-world digital city" (in French) – one thing we have seen over and over is that urban innovation initiatives need cities at the helm to guide processes, ask big questions, open things up to debate, and imagine new models of multi-actor governance. What we are seeing is "a

new role for regional authorities" as "mediator and driver" of systemic change at the territorial level.

Self Data is based on the idea of an innovation that is not solely focused on disruption, or on a startup with a "revolutionary" new offer. Self Data asks adopters to bet on the future of their cities. True, new services will inevitably emerge from any ambitious regional Self Data program. But more importantly, taking this direction involves radically transforming the way we look at personal data, and creating a more distributed, multi-purpose, value-enabling ecosystem that gets data holders, citizens, social enterprises, researchers and entrepreneurs (public and general interest) involved in the dynamic. Who better than local authorities to guide regional stakeholders into this paradigm shift?

The local public actor must play the role of Self Data driver, mediator and facilitator of a new model of personal data governance. **A metropolitan authority is unparalleled in its capacity to make Self Data a reality for three reasons:**

1) As a data collector, it can lead by example, and set up sharing protocols to return the data it holds about its citizens to them. This adds legitimacy to a call to ask others to take the same approach, starting with its own digital service providers. Local government has both the necessary means and the action force to convince regional orga-

nizations to set up their own data transmission channels;

2) By driving data innovation, cities can foster new forms of personal data governance by a) promoting and piloting models and projects that enable citizens to reuse their data via frameworks of trust, and b) getting a variety of private and public actors and members of civil society around the same table and offering them the space to experiment together.

3) As a key relational intermediary, in its capacity as first point of contact for all things citizen related, the regional authorities are in the ideal position to establish a more symmetrical, active role to individuals. Citizens are presently enriching urban services with their data, rather than their own lives. By utilizing digital mediation and co-design, they can work with citizens to imagine Self Data scenarios that speak to them, and create use cases that not only correspond to citizens' needs, but also to the challenges facing the wider region as a whole.

We could not escape the conclusion that municipal actors have a special role to play in making Self Data a reality. Our discussions with the French cities of Greater Lyon (already a MesInfos pilot study partner), the City of Nantes and La Rochelle confirmed what we had concluded. This led to our launching the Self Data Cities project

(September 2018 to 2019) with these three cities as our partners in the field. We began by defining the role of each regional authority in themed experiments formulated with citizen empowerment in mind. The themes in Nantes Métropole were "understand and limit the impact of my food choices", "calculate and reduce my home's carbon footprint", and "contribute to the production of renewable energies in my neighborhood/city". In La Rochelle, the themes were "better manage my mobility budget" and "contribute to improving the transportation offering in my region". And for Greater Lyon, the themes were "oversee my social services rights and obligations" and "understand my (social welfare) rights and entitlements".

Each territory would be exploring a different approach to Self Data, and so to each its own use cases and challenges, according to the agenda each was going to pursue. The choices that would inform project implementation – actors' roles, positioning – would be crucial, because they could lead to new and unique data sharing models.

SHARING PERSONAL DATA: GOVERNANCE MODELS

"The publications, studies, research programmes, and think tank agendas evincing interest the smart city are legion But protecting personal data is still very much the poor relation of these efforts. Individuals remain, at worst, problems that need solving – just as they were in smart city 1.0. At best, for some promoters of participative and contributive cities at least, they are treated like mobile smartphones whose data is essential to the proper conduct of urban affairs."

La Plateforme d'une ville. LINC - CNIL (in French; project summary, in English)

All too often we have attended conferences and workshops devoted to finding answers to questions like "How do we get people to share their data with us?" or, the slightly better, "Which legal and technical data sharing approaches will also protect people's privacy?"

Rarely – far too rarely – have we heard questions like, "Why share personal data with the individuals who generate them? To what end?" We have been treating personal data as if it were stolen booty to be parceled out in secret, or a hostage to protect during the never-ending

struggle between proponents of privacy protection on the one hand, and economic gain on the other.

The answer to the question of "who is allowed to do things with personal data" – i.e., who has the privilege of deriving value from personal data – is important, absolutely. It was our point of departure for the MesInfos research: how to enable individuals to personally derive utility from their data. But the question "who" necessarily entails reflection about the purpose – "to what end" – because it is in the realm of purpose that the risks and abuses reside.

So are these two strategies mutually exclusive? Or can we remove the "tech goggles, which falsely suggest a binary trade-off between privacy and innovation" that author Ben Green describes in his book, *The Smart Enough City?* A happy medium is what Self Data points toward: a framework that puts individuals in control of their personal data, and at the same time allows for new value-creating uses to emerge in response to individual, collective and civil social challenges.

WHAT IS THE BEST WAY TO SHARE PERSONAL DATA? 5 SELF DATA GOVERNANCE MODELS

During our work on MesInfos, we explored and gained experience with a specific type of data sharing model: the personal cloud. While it does offer many advantages, we believe that no single model is uniquely capable of rendering individuals masters over their data. There are at least five "off the shelf" models, including personal clouds, all of which can be hybridized, modified, etc. Given that there isn't a one-size-fits-all solution that can unilaterally support the potential of Self Data, the role of the local public entity, leader of the movement, will be to guide collaborators towards one or another personal data sharing governance model.

The governance models we describe here are of differing orders of magnitude. The first two are very different from a technical standpoint (direct transfer, personal cloud), which is not to say that they are neutral – everything depends on how they are used, who drives their use, etc. The other three models – trusted third party platform, data co-operative, and civic data trust – are organizing principles that can operate in tandem with the first two. By providing detailed explanations of each of these models of governance, we seek to inspire metropolitan public

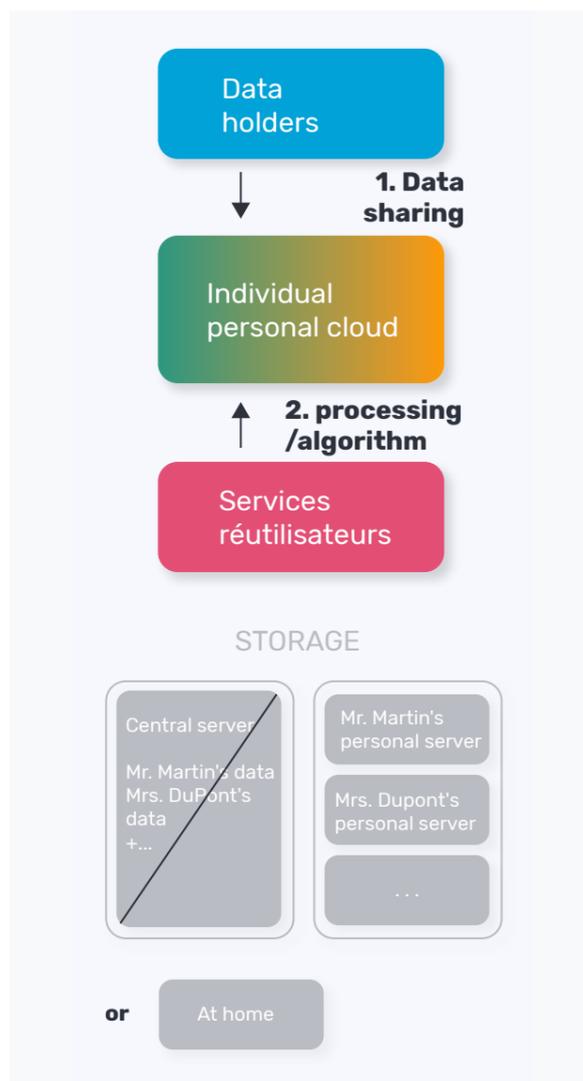
actors to act, and equip them with the tools they need to select among them.

TECHNICALLY SPEAKING, TWO VERY DIFFERENT MODELS: THE PERSONAL CLOUD VS. DIRECT TRANSFER

1) The personal cloud

The personal cloud is like a digital home: a space where individuals can aggregate their data from different sources on their personal server, not one run by an organization. The advantage of having a personal server is that any services designed to help them derive value from their data will run on that server – a process known as on-board computing – the data don't leave the server. **With a personal cloud, the services come to the data, and the data stays put.** Such services can be developed by anyone based on synthetic datasets, which are representative (in form), but not actual (in detail). No data needs to leave their digital home, and yet individuals can still enjoy services and apps that mobilize and triangulate their data that they find on their cloud's "Store". Individuals have complete control over their data, which is stored on their machine, virtually or locally (self-hosting is also a possibility).

Personal clouds are still very new. Low adoption by individuals and data holders alike is a barrier



to their development. Even though there are many advantages to using them, for data holders personal clouds are yet another intermediary between

them and their users and clients, plus there are transmission channels and connectors to keep up to date. . . For individuals, they are another tool to master. For reusers – the ones who develop third-party services – personal cloud projects require extra investment: to create trustworthy services, they have to learn to work with data-holding organizations' data controllers, and therefore align their efforts with other actors whose data management approaches may differ from theirs. They also have to adapt technically to the operating system of every personal cloud platform they encounter. Most developers are used to creating services for well-established (GAFA) operating systems – Apple's iOS or Google's Android, to name the two biggest – or within the massive, self-contained ecosystems like Facebook. This is a difficult adjustment to make when the technology is just emerging and has very few users.

In the digital economy, where personal data sharing with third-party services (personal organizers, public transport apps, etc.) is integral to their functioning, here we consider the advantages of using dedicated data-sharing environments such as personal clouds. Data crossing takes place in a personal digital space, carried out by a trusted third party provider as host. Even though for the moment, the personal cloud sector is largely focused on individuals: "a personal cloud = an individual", research is underway looking at how they can be used to derive collective use from perso-

nal data, for example by allowing algorithms to run over thousands of clouds without the data leaving those clouds. This means personal data could be used to power big data analytics and still remain under the control of the individual they concern. In addition, personal cloud sharing capabilities (data sharing from individual to individual, data pooling, etc.) are improving.

Exemples

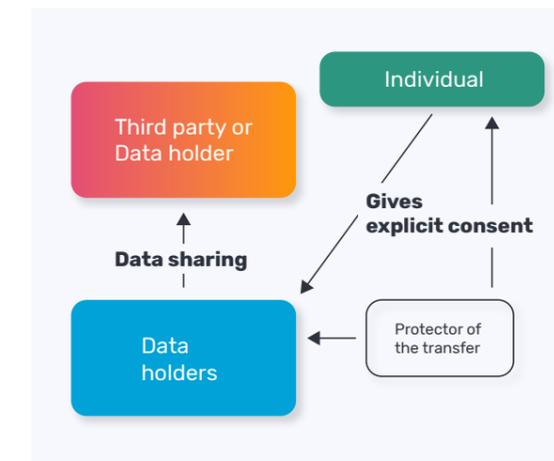


2) Direct transfer

Direct transfer is grounded on a key principle: consent. Data sharing takes place directly between data controllers with the explicit, revocable consent of individuals, for services provision or so they can participate in public interest projects or research projects. For example, energy provider Enedis created the "Enedis Data Connect" project. Customers who have a Linky smart meter are offered services that will connect, with their consent, to Enedis' information system and help them derive use from their consumption data. This model has a great advantage: fewer actors, and therefore less complexity as regards implementation. A contract can potentially be signed between the data holder and the reuse service to specify processes, levels of data protection, usage of the bearer's infrastructure to retrieve data (volume, regularity, etc.). Some services could potentially be "blacklisted" and hence not allowed to connect to the data holder's system for security reasons, for example.

Although there are few organizations using this kind of transfer protocol – the model we describe here doesn't really exist in the minds of actors outside the digital – the likelihood that it will inspire uptake is probably quite high, because it does not "mess with" coding, and is consistent with the current digital economy.

Everyone's responsibilities are clearly defined, and, above all, it represents a mitigated legal risk for both holders and re-users.



And yet it does not really allow individuals to get a 360° overview of their data, to really master it: third-party services are not going to connect to 1000 different holders' APIs and sign 1000 different contracts! This means that a sector-based approach might be more appropriate, and that data reuse potential, from the perspective of the reusers, might be limited somewhat. Another drawback is that individuals are "merely" consent-givers. Even though consent is informed and explicit, individuals can't see the data "moving", nor can they reuse their data themselves. Also, the process is fragmented: consent must be given to two actors, minimum, for each service. Moreover, unlike the personal

cloud, the data flows to the services, which means that personal data are duplicated and stored within each service. This contributes to the proliferation of personal data, and makes personal privacy mechanically vulnerable.

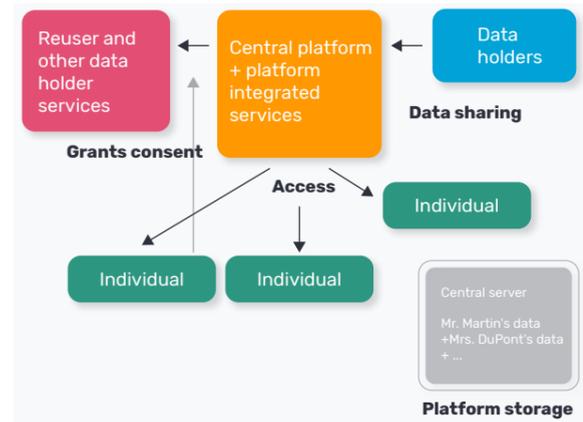
A separate entity can, however, play the role of trusted third party between individuals and the other two organizations (data holders and reusers). We think of them as "transfer guarantors". They ensure the security and authenticity of data sharing and provide a dashboard for individuals (and organizations) to oversee and manage their data rights and obligations related to portability (and thus consent/sharing), data erasure, modification, etc.

THREE MODELS THAT ARE TECHNICALLY SIMILAR, BUT WHOSE GOVERNANCE APPROACH DIFFERS: THE TRUSTED THIRD PARTY PLATFORM, CIVIC DATA TRUST AND DATA COOPERATIVE.

Exemples



3) The "third party platform"



A key facet of this model is basically that it is the 2.0 version of a digital vault: a portal platform/personal space where users can retrieve and organize documents and data from multiple sources, share them and also take advantage of services from third parties or ones that have been integrated into the platform. So how is a personal cloud different? Storage. On a personal cloud, data storage is centralized on a server, and data processing by third parties is performed elsewhere than locally. France's Shared Medical Record is an example of a "trusted third party" platform, although today it supports mostly document rather than data aggregation, and is used chiefly by users to share data with health professionals rather than reuse them via dedicated services.

The keyword in the trusted third party platform concept is "trust" – the model is, after all, very close to Google's – so without it, the current paradigm would not change! The platform operator plays the role of transfer guarantor and also centralizes storage. Control over the tool and the data rests more in the hands of the organization, as provider of the platform and manager of the services it offers. This is similar to the way services such as Amazon and Netflix use platforms (but minus a PIMS) to promote their own products and services ... Individuals still maintain a global view of their data. Nowadays, these platforms are often provided by large institutions. Data reuse tends to be focused on integrated services and personal data sharing (eg: France's DigiPoste postal platform allows users to organize the documents and data needed to build a real estate application file and share it to an agency).

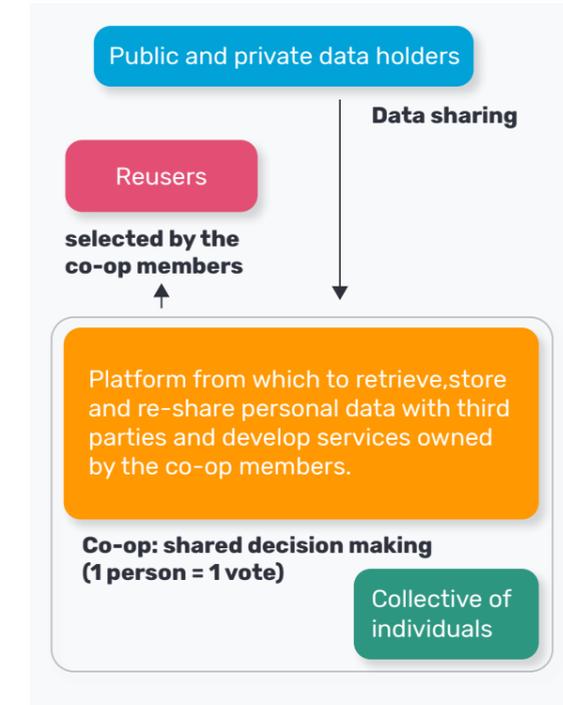
Exemples



4) The data co-operative

What if individuals formed a group to jointly manage their data, and then decided how they wanted to use them and share them, collectively? This is the route of the (still rare) data co-operative. Based on the principle that 1 individual=1 vote, data co-op members jointly develop tools and services (chat apps, search engines, etc.) that eventually grant them control over their data from A to Z. The collective might also simply decide together to share them via a platform, for example to contribute to public knowledge creation.

In this model, individuals regain control of how their personal data is used by others (sharing) and of the uses they themselves will make of them (services creation). They might adopt existing open source, free-standing services (email or chat services, etc.) or develop their own. This model is particularly supportive of collective initiatives that require data pooling. In fact, personal data can be made more valuable when not used in isolation. According to researchers Antonio Casilli and Paola Tubaro, "there is nothing more collective than personal data". To this, we add the words of France's National Centre for Scientific Research Associate Scientific Director Lionel Maurel: "any notion of collective data must include a collective decision making power that incontestably and severally resides in the hands of its members".



Even though data collectives open up new horizons in data governance, they are confined by their own set of limitations: first and foremost by the amount of time members must dedicate to governance matters (the bigger the collective, the more complicated this becomes). Although that burden can be restrictive, the critical mass created by large numbers of members must also be weighty enough to flesh out a quality digital service. In addition, in the absence of that critical mass, there is the significantly heightened risk that pooled data shared by the collec-

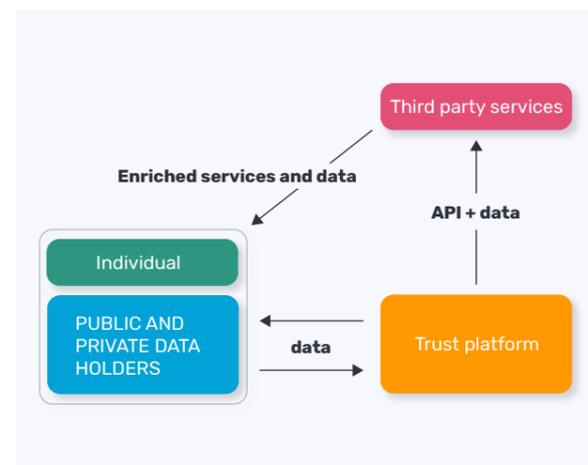
tive can be re-identified to the individual: is data anonymity really possible in a collective? Members cannot see other members' data, but everyone knows the names of the other collective members (via the roster); voting procedures may also be open (debate, joint decision-making). Also, a cooperative is under no obligation to guarantee that data usage will be virtuous (a cooperative has the right to grant its members permission to resell their personal data for a few cents), and there are numerous examples of collectives being as heavy handed with their employees as organizations operating under any classic mode of governance.

Data co-operatives - and their derivatives - remain, however, one of the few models that make it possible to foster collective uses of data and where individuals remain masters of these data.

Exemples



5) The civic data trust



According to the Open Data Institute, a civic data trust is “a legal structure that allows independent management of data by a trusted third party”. The trustee body can comprise public actors, private actors and members of civil society. Its governance rules can be multiple, but those rules must facilitate consensus on the trust’s use of the personal and non-personal data entrusted to it by individuals and data holders.

The data can be physically stored and accessed via a platform (see our diagram), or can be left where they are, and the trust becomes the “trusted guarantor” who oversees private and secure data transfer – for example to contribute to causes that serve the public interest, public policy discus-

sions, etc. – according to the conditions defined by the group. This model also represents a step beyond Open Data. Public data on their own do not really facilitate the emerging services market, especially ones designed to serve the public good, but crossing public data with personal data may contribute to the creation of entirely new uses.

Today, in the majority of the work being done around this model, little emphasis is placed on personal data and the role of individuals. In fact, personal data play only a peripheral role in regional data platform development, data management and the wider domain of civic data trusts. But some trust builders are starting to take an interest in it. Similar to third party platforms, the crucial question for individuals is, “Who is the third party, exactly?” In Toronto, when Google offered to set up a civic data trust, the proposal immediately raised questions such as, “Will Google be the one to choose the trustees given a seat at the table when it comes to defining the rules?” Some are starting to come up with alternatives – one is to entrust the [National Library of Toronto](#) with the responsibility of oversight. As noted by [Sean McDonald](#), “It is just as easy to build many Civic Data Trusts, for example, as it is to build one, so a city could organize them according to use case or representative group or public need. Trusts are structurally easy to replicate, and given data’s ability to support non-competitive re-use, it makes

far more sense to treat them as smaller, agile units than one overarching, governing trust”. In any case, for a civic data trust to work, data holders have to agree to share the data they hold. And in keeping with the spirit of Self Data, individuals must be treated as more than merely the data they contribute – they must be granted real power within the governance framework to define future uses.

Examples



C:RONOS

MAKING A CHOICE: WHY CHOOSE ONE MODEL OF GOVERNANCE OVER ANOTHER?

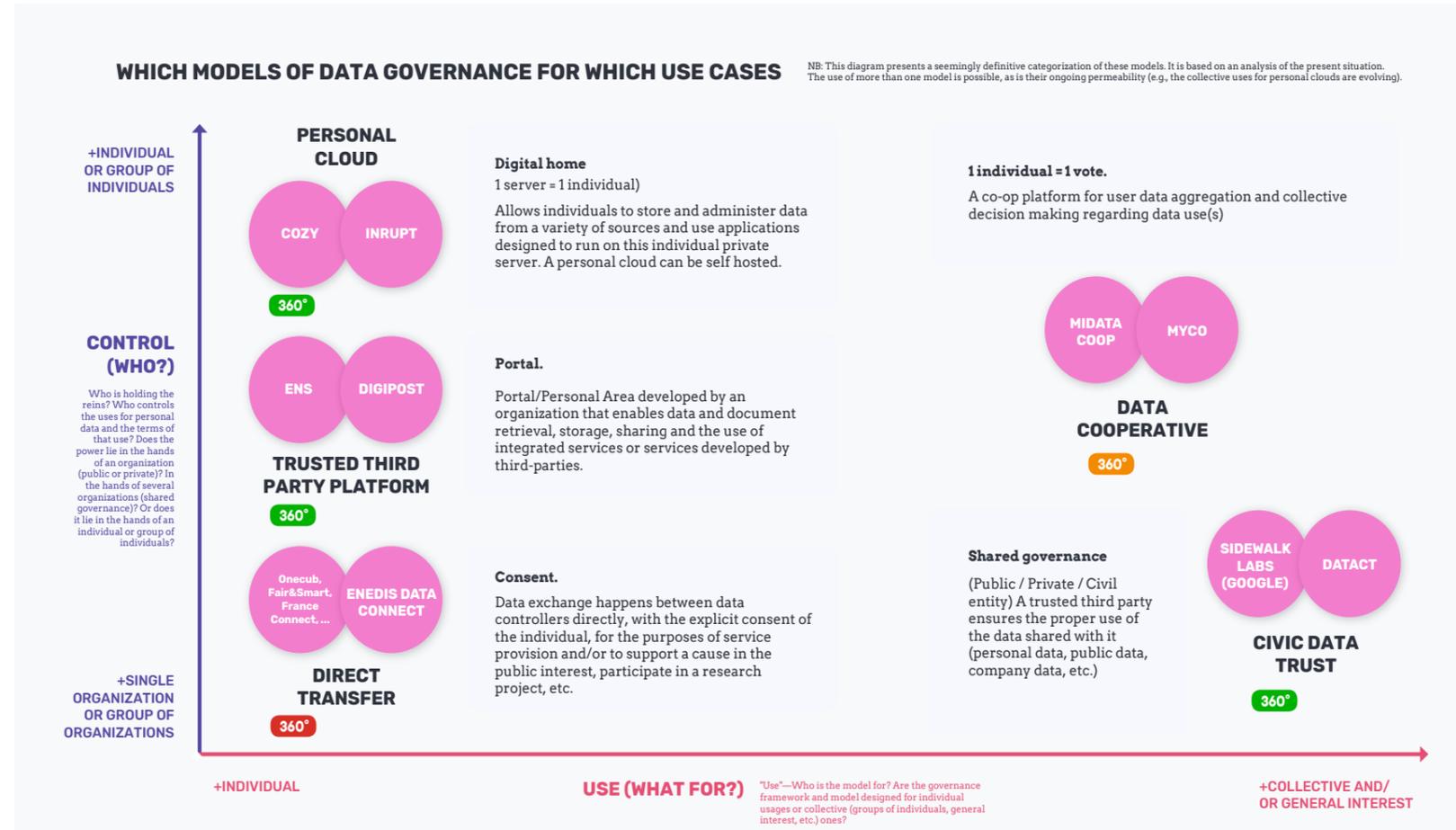
All five of these personal data sharing models have strengths and weaknesses. Local public actors embarking on Self Data will have to make choices. One model is not necessarily more virtuous than another – everything depends on the needs and constraints of the actors who oversee it.

Some may want to set up a Self Data project whose purpose is to create massively collective uses for personal data, while others may wish

to do the opposite, and provide citizens with individualized services. Some may want to let people control their data – and make them entirely responsible for their data – and some may find that the burden is too heavy for citizens, and suggest that they go through a trusted third party.

The choice of one model over another rests on the answers to two key questions:

- » The question of control: who is holding the reins? Does the power over what kinds of personal data to collect and how to use it rest



in the hands of a single (public or private) organization, or is it controlled by several organizations (shared governance), a group of individuals, or the individual alone?

- » The question of use(s): who is the model destined for? Do the data governance and governance framework favor individual uses or collective uses (for groups, to contribute to the common good, etc.)?

By positioning the models of data governance along these two axes, we hope to equip and guide actors making the choice among them. The personal cloud, trusted third party platform, and direct transfer all present mostly individual use possibilities. The personal cloud puts the keys of a digital home, and control of the personal data it contains, in the hands of individuals; for the third party platform, the platform provider organization is in the driver’s seat; and in data transfer, the data-holding organizations have power over data sharing decisions. Both the data co-operative and the civic data trust have vast potential for collective use, but radically different loci of control: the first gives empire over the data to individuals, whereas the second offers them, at best, a marginalized seat at the decision making table.

Despite the utility of having clear definitions, we advise you to remain cautious of the way

they can be enacted. According to the above, for example, Facebook could be classed as a trusted third party platform provider – and so could Microsoft, as evidenced with its recent launch of BALI. Careful consideration of the data storage aspect is equally paramount. If a data co-operative stores data in a country where anyone can access it, individual control over the data will take a hit. These are essentially the same issues that pertain to any platform in the digital economy. Everything depends on who oversees the model, and according to which rules. Then another question arises: what role should you, the public actor, play in the model? Tools and services provider? Public tender sponsor? And who ensures the framework put in place respects every stakeholder involved? No choice is “wrong”, and whatever course you take will be experimental in some way, given the emerging nature of the domain of Self Data.

Then there is the economic landscape to consider. If governance is by third party, who is to say that investors will not eventually force them to act like any other service, and hand over users’ personal data? Lionel Maurel (Calimag, in French) asked this very question after Tim Berners-Lee announced that his startup Inrupt had secured funding from venture capital giants Glasswing. “If Solid meets with success (and here’s to hoping it does), what guarantee do we have that Inrupt won’t suddenly change its

model. . . . The risk of a buyout by one of the web’s Leviathans like Apple, Google, or Facebook should not be dismissed as sheer fantasy”.

So what will reassure users and build the necessary trust? Are the use of open source technologies and the possibility of taking one’s data elsewhere if the service is bought back sufficient guarantees? Maurel invites us to consider the possibility that the relationship can be a virtuous one along the lines of Coopcycle, a European federation of bike messenger co-ops “my current thinking is that we need to consider converging the digital Commons and the solidarity economy to develop models that are part of the economic sphere, but also aim to limit profitability as an endgame from the very start, and reinvest market and other actors’ efforts into serving the public interest”.

An aerial photograph of a dense city, likely Tokyo, with the Tokyo Tower visible in the center. The image is overlaid with a purple-to-blue gradient. Two large, semi-transparent circles are positioned over the city. The number '02' is written in a large, white, outlined font, centered over the right circle.

02

**“PERSONAL DATA AND
CITIES” – AN OVERVIEW
OF INITIATIVES AROUND
THE WORLD**

02

“PERSONAL DATA AND CITIES” – AN OVERVIEW OF INITIATIVES AROUND THE WORLD

CITIES AND PERSONAL DATA: WHAT WAYS AHEAD FOR CITIZENS?

After several months researching the Self Data Cities-style initiatives happening around the world, we understand that our definition what Self Data is and does is still fairly advanced. And yet, some cities and local authorities have begun to see the value – at the crossroads of citizen empowerment and technological development – that giving citizens control over their data represents.

For example, the Cities Coalition for Digital Rights – a joint initiative launched by the cities of Amsterdam, Barcelona and New York, with the support of UN-Habitat, EURO CITIES, and the UCLG – worked in conjunction with a number of cities worldwide to help them protect, promote, and monitor residents’ and visitors’ digital rights. Personal data has clearly become an increasingly essential resource for cities, both in terms of citywide operations and public services improvement. Sometimes citizens directly provide those data, but public and regional actors are typically the data holders. The Smart City was practically unavoidable as a topic, including its implications for personal data management and use, individual privacy, and the introduction of new technologies into the public sphere. Smart City critiques and

analyses abound; we will not expand further on the topic here. However, we invite you to consult the CHILD’s report “The platform of a city: Personal data as the cornerstone of the smart city” (in French; partial English translation) which outlines not only the stakes involved, but also the pathways towards reconciling “the datafication of the city and the protection of civil liberties”.

Today, when the words “personal data sharing” emerge in context of cities, very rarely do they refer to Self Data autonomy or to frameworks that put individuals in control of data exchange. Instead, the “sharing” takes place between a variety of public and private actors. The Open Data Institute in London has identified a number of data-sharing transactions between the private and public sectors, and classified them according to two criteria: public actor role and objective pursued. The public entity can act as user, collaborator, customer, financier or regulator, and seek to stimulate innovation, promote competition, improve transparency and accountability over data, support research and political planning, support operational decision making, or assume regulatory functions. The exchange of data between cities and companies is increasing, either in the form of collaborations or in the form of a legal obligation towards companies.

There are many examples of cities cooperating with private actors in the domain of mobi-

lity. There is the example of Vodafone, who is working with the city of Barcelona to enrich its Open Data Portal; the data sharing agreement between Transport for London and Waze, which gave rise to WiFi connectivity across the underground plus the enrichment of Waze services using TfL crowdsourced data; and across the pond, in San Francisco, a collab with Uber and RapidSOS intends to improve the rescue data provided to the city’s emergency services. New York City, on the other hand, has taken the route of legal injunction against certain companies to obtain real estate rental data from Airbnb, for example, and traffic data generated by taxi companies Uber and Lyft. There are also platforms like Strava Metro or Uber Movement that grant cities (but not only) access to mobility data. On Strava Metro, datasets are enriched by swathes of users via their personal account, while Uber Movement uses its customers’ and drivers’ anonymized data.

Today, Self Data is emerging as a complementary or even alternative to the existing data recovery system – by including people in the decisions made about data sharing, it turns them into masters of their data.

In compiling this cities and personal data project overview, we found it useful to categorize each approach in terms of how it contributes to achieving the overarching, and as yet unrea-

lized, Self Data concept. We put the approaches into 4 categories:

- 1) individual privacy;
- 2) citizen decision making;
- 3) metropolitan data generation using citizen sensors;
- 4) “pure” Self Data.

None of these categories are exclusive – several projects could belong to all four. We found that few cities are considering the possibility of actually including citizens in the governance of personal data. The majority deal with privacy or with Open Data (a subject we do not discuss here, as it is outside the scope of this study). Even though cities are trying to build new infrastructures, none are looking to establish a Self Data dynamic, which would require specific measures to address both data sharing from holders to individuals and overall transparency.

One thing to note is that the majority of the initiatives we found are being conducted in Western countries. For this review, we drew largely from reports and articles published by major European and North American institutions and think tanks including the Open Data Institute, Nesta, CNIL, etc. But personal data sharing and Mydata/Self Data is pervasive: there are almost twenty MyData hubs up and running across the globe, and although they can be found mostly

in Europe, the hubs in Cameroon, Japan, Brazil, and Korea are particularly active and will certainly launch initiatives in these territories.

PROTECTING CITIZENS’ PRIVACY

Examples of measures taken to protect individuals’ privacy were the most prevalent in our research. In the Self Data dynamic, data privacy and security are prerequisites when establishing personal data recovery, storage, and management channels..

The number of cities asking themselves about the citizen-generated personal data they possess – especially how that data are managed and used – is significant, and growing. Issues abound relative to the lack of transparency about which data are being held, protecting citizens’ privacy, and the proliferation of data-capturing systems in public places.

Before devising strategies for ways to share personal data with a city’s inhabitants, some public actors have raised questions about the infrastructure needed to capture data and use it ethically. In cities where sensors are more and more numerous, and where Smart City initiatives have taken off, it is essential to think about how these devices work and how to protect people’s privacy. A number of cities have

taken steps in this direction, to varying degrees.

We also note that many of these cities have widened the frontiers of Open Data in recent years, and created platforms that facilitate city data sharing and reuse by third parties. Is the terrain of open data sharing the right place to begin to consider the protection of citizens' data privacy? After its Open Data campaign in 2015, the City of New York produced its [Guidelines for the Internet of Things](#), a series of recommendations designed to help others protect citizens' privacy. Echoing that approach, Seattle produced a [Privacy Principles](#) toolkit to help city departments integrate data transparency and other Open Data-related guidelines. The publication was issued after the city's installation of a massive network of sensors in the city and deployment of drones was met with serious backlash in 2013. Finally, there is the example of San José, California, whose [Digital Privacy Working Group](#) (launched in May 2018) is actively meeting once a month.

There are many other examples of such initiatives. Nantes Métropole has a Metropolitan Data Charter; and Boston, Montreal, Amsterdam, and Barcelona have also taken an explicit stance regarding privacy, a cornerstone of Self Data.

INCLUDING CITIZENS IN DECISION MAKING

Our second category takes a step back and questions how approaches to personal data governance might include citizens in metropolitan decision making. This might take place across a number of mediums, and provided that participants' privacy is respected, these tools can be essential bridges between a metropolitan area and its inhabitants. Several platforms have been developed that give citizens a voice while leaving them in control of their data.

Emblematic of this category — although still an isolated example — the [Decidim](#) open source digital platform supports citizen participation that extends beyond the walls of any city. It is being tested in 35 cities, including Barcelona, Roubaix, Helsinki, Mexico, and Waterloo. The Decidim platform — available for use by private as well as public organizations such as city councils, social enterprises, universities, NGOs, unions, neighborhood collectives, co-ops — allows users to configure participation areas or “spaces” to launch initiatives, hold assemblies, conduct processes, or handle consultations. . . and enrich all of these offerings via support for face-to-face meetings, surveys, proposals, votes, follow-up reports, and comments. . . all within the parameters of an integrated module enabling individuals to control authentication data transmission, anonymity should

they sign a petition, or data sharing via dashboards (see for example the [BCNOW](#) dashboard) designed for collective use.

Participatory democracy platforms are cropping up all over, in places like Turin, ([decidiTorino](#)), Sydney ([Sydney Your Say](#)), and Madrid ([DecideMadrid](#)). Over time, the forging of stronger ties between such “civic tech” and Self Data — along the lines of Decidim — is highly likely. Imagine if citizens could choose to pool their data to push for specific measures to be taken (eg: residents living near a factory pooling their health data)? Collectives would also be able to use a platform to collaboratively develop Self Data services that serve public policy objectives, like a CO2 reduction app for members, or to publish a call to assert their right to share their personal data with public actors to improve mobility plans in their region...

USING DEDICATED SENSORS TO GENERATE CITIZEN DATA

A number of citizen sensing initiatives — equipping citizens with data sensors that they control — have emerged, providing citizens with the tools to capture and share data the city can use to improve its offerings. This approach is closer to what we consider Self Data, in that citizens produce data that they can choose to share. However, the data are often generated solely to serve a city's

purposes, while their potential for personal use is rarely fully developed (one limited example is the French initiative [Citoyens Capteurs](#)). Citizen sensors are tools for citizen emancipation — they grant citizens access to technical tools they can use to “measure” their city and in doing so, acquire a policy-oriented awareness of pollution or radioactivity, for example. But they are also extremely useful and also free sources of data that cities would not otherwise be able to access, especially measurements taken in private spaces. Sensors should not become tools that cities use to manipulate their citizens — they should offer concrete opportunities to integrate inhabitants in metropolitan operations management and act as tools for generating shared knowledge.

The [Making Sense](#) project — a representative example of a citizen sensing experiment that gives inhabitants control over how their data is used — is on the right track, by offering workshops to help citizen participants learn to measure radioactivity levels in Belgium and the Netherlands (see below for more on this). Another initiative called [Citizen Sense](#) is in the same vein. It contextualizes, questions, and develops avenues for democratized environmental action through citizen sensing practices. Data detection is used to serve three complementary purposes: measuring pollution, observing flora and fauna, and promoting sustainable

development in urban areas. One way is to examine how vegetation can improve or respond to air quality via its [phyto-sensor toolkit](#). Let us therefore make a clear distinction between sensor initiatives in cities that people have little to no control over (eg: digital advertising screens in public transport that track audience views) and initiatives where Individuals not only have control over the data they collect, but are also active participants in its use, for example by opting to share it to serve the interests of their communities and the public good more broadly.

IMPLEMENTING SELF DATA

Our last category presents what we call 100% “pure” Self Data experiments, in that here, individuals personally retrieve their data and do whatever they choose with them. Such initiatives are few and far between, and are a municipal rarity when compared with the number of projects that focus on privacy or citizen sensing. It is mostly at the national level that the Self Data approach is showing the most promise, with different versions emerging in a variety of contexts. Japan is developing its [Data Banks \(or Information Banks\)](#) — databases for the secure collection, management, and sharing of personal data held by individuals or companies with the explicit consent of the data subject. Individuals will be able to choose the companies

they wish to share data with via the information bank system, which will transmit the data to those organizations under conditions that have been agreed upon in advance. An experiment was launched in September 2018 by DAC Consortium Inc., in partnership with Hitachi. Needs have also already been identified in the field of tourism, health, and agriculture. And Japan's NTT Data, one of Fling's global partners, will soon be launching a [MesInfos](#) research project in tandem with partners from the banking, insurance, and retail sectors.

Over in Finland, the government agency responsible for promoting safety in the Finnish Trafi transport system has begun a [MyData pilot](#) to catalyze the transformation of its personal data management system. The agency is an important driving data holder, gathering resources from drivers licenses and driving professionals across Finland and Sweden. The pilot's aim was to enable them to offer these data subjects improved data access and control, and significantly, the real ability for them to reuse data. Until now, the agency's model was based on obtaining subjects' consent to the flow of their data between companies and organizations. Based on MyData principles, the pilot focused on a use case involving 200 testers. Trafi developed a MyData Wallet API and personal platform to enable professional drivers to share, manage, and make use of their driving right and professional

qualification data to demonstrate their eligibility for specific jobs and streamline procedures associated with their work.

Another good example of a Self Data initiative is Digime, a project positioned in the health sector in Iceland. Digime is a personal data aggregation and management platform. The company managed to set up an experiment in partnership with the government and its Ministry of Health. Testers were able to retrieve personal health data stored in the state's information system. Beyond data retrieval, Digime seeks to encourage the development of dedicated applications that foster new uses of the data (to track treatment, streamline sharing with caregivers and health professionals, etc.).

At the city level, we draw your attention to projects like Milan's *Digital Citizen Folder*, a private and secure digital repository where personal documents are stored for all citizens. From a single point they can access all their personal information and data directly from the municipality's web portal. The cities of Amsterdam, with its *Digital Register* project, Trento and its *My Data Store* platform, and Toronto and London's consideration of the *Civic Data Trust* are each venturing into the Self Data ecosystem; we will shine a spotlight on their projects in the sections to follow. The French approach, outlined in later chapters of this report, places individuals at the heart of their data. Our groundbreaking

stance on personal data and its use by and for individuals is still unique, as demonstrated by city initiatives in Rennes and Grenoble, and of course put into practice within the *Self Data Cities* projects we have led in conjunction with the cities of Nantes Métropole, La Rochelle, and Greater Lyon.

SPOTLIGHT: A CITY-BY-CITY PRESENTATION OF PERSONAL DATA INITIATIVES

Each of the following examples sheds light on what we consider to be a major metropolitan initiative. Rather than being standalone projects, however, each is woven into the fabric of a (typically European) initiative that involves several cities. Those initiatives — including DECODE, Making Sense, Solid/Inrupt, Sharing Cities, Eurocities, etc — are building the bridges that lead to even more widespread sharing, communication, and collaboration.

PROJECT CATEGORIES:

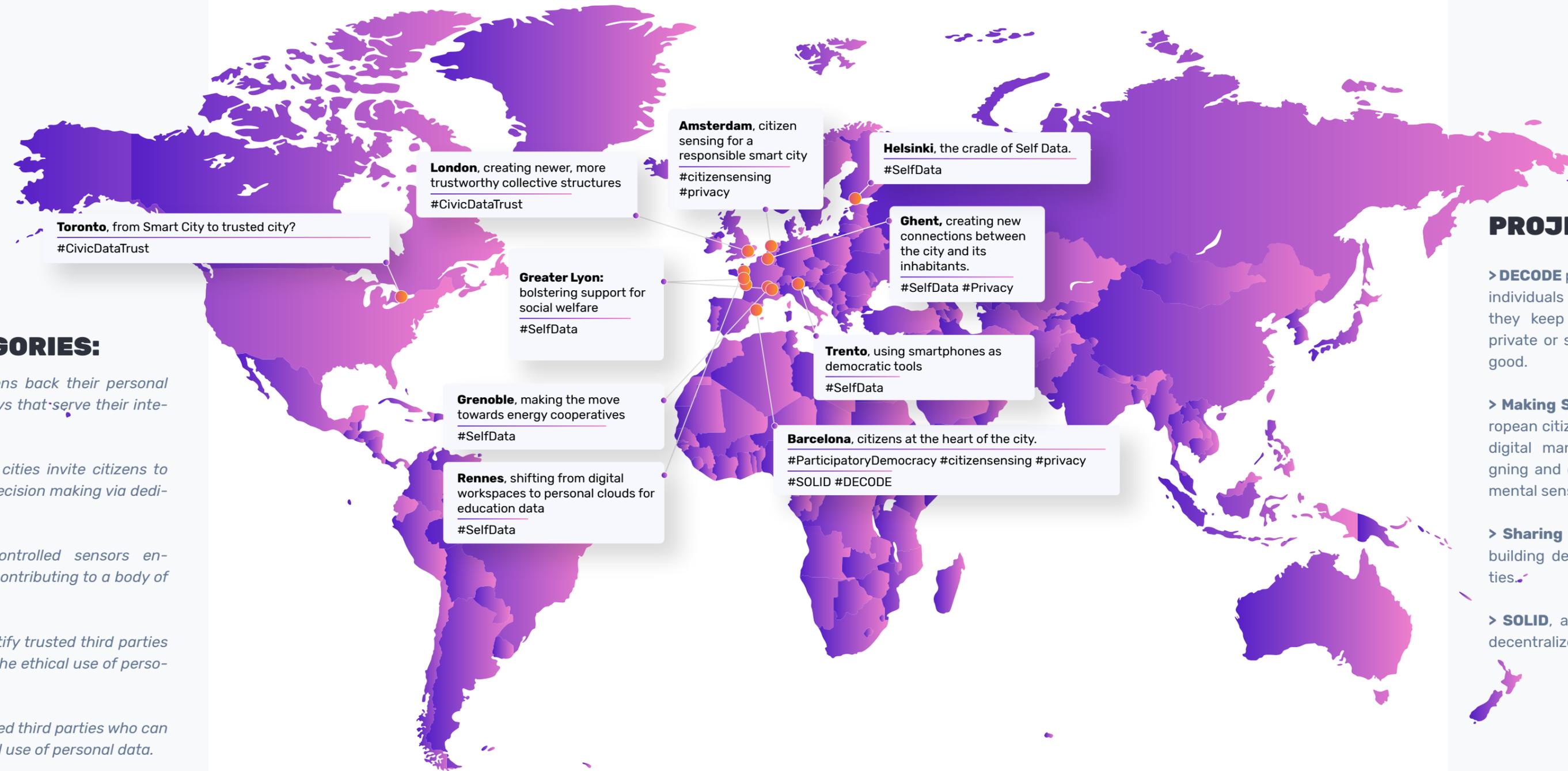
#SelfData : cities give citizens back their personal data so they can use it in ways that serve their interests and the common good.

#ParticipatoryDemocracy : cities invite citizens to actively participate in policy decision making via dedicated platforms.

#citizensensing : citizen-controlled sensors enable them to help the city by contributing to a body of shared knowledge.

#CivicDataTrust : cities identify trusted third parties who can oversee and ensure the ethical use of personal data.

#Privacy : cities identify trusted third parties who can oversee and ensure the ethical use of personal data.



PROJECTS:

> **DECODE** provides tools that put individuals in control of whether they keep their personal data private or share it for the public good.

> **Making Sense**, empowers European citizens through personal digital manufacturing, co-designing and deploying of environmental sensors.

> **Sharing Cities**, works toward building democratized smart cities.

> **SOLID**, a new paradigm for a decentralized Web.

BARCELONA, CITIZENS AT THE HEART OF THE CITY

Barcelona has been a pioneer Smart City city since 2011. It has 122 projects organized into 22 programs. When Mayor Ada Colau was elected in 2015, the city began to move in a different direction, and make citizens a central part of its metropolitan project development. Ethical concerns were the impetus for the city's growing interest in open standards, Data Commons, and interoperability.

Since then, a variety of projects have emerged that focus on the roles public actors might play in personal data management. Two large-scale projects have attracted the city's attention: the European *DECODE project*, which has two pilots underway in the city; and the earlier *Making Sense project* (2015-2017) which focused on citizen sensing implementation. The city places a high value on the sensing approach, which offers citizens the opportunity to create data about the city and control how it is used. Two projects are notable: *Sentilo*, which established an open source sensor and data collection platform; and *CityOS*, which deployed another open platform for data analysis. As part of *DECODE*, the city built their *Barcelona Now Platform*, which brings a range of city data together in one place, including data gathered by citizens. Finally, *the Citizen Science Data Governance pilot* (DECODE) gives inhabitants the opportunity

to install and control sensors that measure different levels of pollution – including air and noise – in their neighborhoods and homes.

The city also pays particular attention to the privacy of its citizens, and this means protecting their personal data. The *Digital Democracy and Data Commons* pilot, which took place between October 2018 and April 2019, enabled the implementation of a DECODE module on the city's Decidim platform, which allowed citizens to participate anonymously in platform-related projects in exchange for a minimal amount of their data.

The DECODE Project (DEcentralised Citizen-owned Data Ecosystems)

DECODE is a 3-year project – January 2017 to December 2019 – funded by the European Commission under the auspices of Horizon 2020. Its objective is to develop practical alternatives to the way people use the Internet today, especially how they use personal data. The goal is to demonstrate the social value inherent in allowing citizens to control their personal data and illustrate different ways

to share them. Now, the question is: how to create a digital economy that is grounded on data generated, controlled, and shared collectively, and that also takes individual privacy into account?

DECODE is not a platform per se – it is a project designed to foster the emergence of open technologies allowing individuals to control and share their data however they see fit. Four pilots were set up in the cities of Barcelona and Amsterdam between 2017 and 2019, to test new technologies developed during the project. Three themes were the focus: the Internet of Things (IoT), open democracy, and the collaborative economy. The issues of trust and privacy were and are central to the solutions being implemented.

AMSTERDAM, CITIZEN SENSING FOR A RESPONSIBLE SMART CITY

The city of Amsterdam has launched a number of personal data-related projects. Two DECODE pilots have been part of the city's agenda since 2017, and will continue until December 2020. One is the *Digital Register* pilot, which allows residents of the city to retrieve personal data stored in municipal databases and use them on their own terms, thanks to its use of the Attribute Based Credentials (ABC) authentication mechanism. The pilot *Gebiedonline* (Neighborhood Online) uses the same system to circumvent the need to provide Facebook credentials when seeking access to local social networks.

In addition, DECODE helped the city to launch an initiative focusing on vacation rental owners' personal data (eg: Airbnb). A recent law requires landlords to furnish data about the rental periods for their property/ies, which must not exceed 30 days per year. Until now, homeowners had to supply a massive amount of information to complete their profiles. The city wanted this process to be more ethical, and by opting to use an Attribute Based Credentials mechanism, a minimal amount of data could be shared by explicit consent.

Upstream of concrete experiments, the *Tada Manifesto* – written and published by professionals from the Amsterdam region in 2017 – invites organizations, companies, and government authorities to read and affirm a set of 6 ethical principles related to cities' responsible use of data.

Amsterdam's drive to ensure transparency in its operations is further underlined by its *City Data* (Dutch only) open data portal. A vast array of datasets have been available to anyone on the platform since 2015, with categories ranging from public spaces to buildings and plots, traffic, health care, the environment, quality of life, permits, subsidies, and more. However, some of the data stored on the portal were generated by sensors that belong to various organizations, and their diverse formats makes reading difficult. *The Sensor Register*, an open registry of all IoTs installed in the city, is seeking to address citizens' lack of control over these surveillance devices, and create more transparency. Unfortunately, their progress has been hindered by an inability to obtain precise information about the devices. Answers to questions about the sensors' locations, and which kinds of data they harvest, are to be found within what are currently siloed departments inside various organizations and businesses.

Making Sense

The Making Sense project (2015-17) was co-funded by the European Commission. Its aims were to explore ways that citizens could appropriate and utilize open source tools allowing them to capture, analyze, and transmit localized environmental data, and in doing so, participate in solving today's pressing environmental issues.

Project efforts were divided between developing the *Smart Citizen Kit* and galvanizing citizens' initiative to install and use the sensors. Over those two years, other European communities and cities explored initiatives in a similar vein: *UrbanAirQ* (Amsterdam) mapped out pollution in the city's streets, *Gamma Sense* (Amsterdam) measured radiation, *Plaça Del Sol* (Barcelona) charted neighborhood noise pollution, *Fab Kids Lab* (Barcelona) and *Smart Kids Lab* (Amsterdam) raised families' awareness about what sensors can do, and *Air Pollution in Prishtina* (Kosovo) sought to get the city's youth involved in resolving the issue of air pollution.

Several concrete outcomes, including the “Citizen Sensing Toolkit,” were delivered.

GHENT, CREATING NEW CONNECTIONS BETWEEN THE CITY AND ITS INHABITANTS

The city of Ghent was inspired by Tim Berners-Lee and his intention to preserve the Internet as a tool for democracy. Their efforts are bolstered by their collaboration with Prof. Ruben Verborgh from the Multimedialab at the University of Ghent.

Belgium has a complex system of governance; people place the majority of their trust in their local authorities when it comes to managing their data and providing trustworthy data authentication protocols and services. City authorities must therefore work in close collaboration with other levels of government as well as with citizens. An important part of this is designing more ethical personal data management systems. Bart Rousseau, Ghent’s Chief Data Officer, has stated that data governance entities may act as regulators, users, creators, and validators of personal data. As part of its *City of People* (Flemish language only) strategy, the city of Ghent wants to empower its citizens by giving them access to technologies they own and control.

The *Mijn Gent* (Flemish language) project wants to give citizens a single profile page they can use to manage their administrative interactions (consult their appointment schedule, sign children up for sports activities, track the status of aid applica-

tions), that will also enable them to fill out forms faster (thanks to the personal information stored on their profile) and contribute to citywide development via an online participation platform. Three features already exist: Ghent Library Services, Ghent Sports Service, and childcare request service.

In 2018, Ghent launched its *Hallo.gent* initiative aiming to provide citizens with their own online domains and personal websites. The system will be able to register citizens for any municipal service and retrieve only the necessary data (proving their entitlement) from the individual’s personal Mijn Gent portal, without the municipality ever needing to collect or store additional personal information. These 100% unique personal websites are similar in design to the PODs designed by Tim Berners-Lee as part of SOLID. The websites can generate and complete applications, which allow individuals to control how their personal data is accessed and shared with local authorities.

SOLID : Linked Data

Tim Berners-Lee, one of the creators of the World Wide Web, wants to (re) decentralize the Web. His idea puts countless existing business models based on harvesting and exploiting users’ personal data into question. What would happen if people were the ones storing their personal data, and hence rendering them inaccessible to any external entity (government, businesses, etc.) without their consent? The web’s heavy hitters (eg: Google, Apple, Facebook, Amazon, Microsoft, the so-called GAFAM group) likely see little to no advantage to adopting this system – but emerging business concerns and local public entities (like city authorities) could very well take up the gauntlet thrown down by Berners-Lee. Ghent is a case in point.

The *Solid platform* could become a springboard for cities looking to reshape their personal data policies. Solid empowers individuals by giving them control over their data. Its design is grounded on a set of conven-

tions and tools for decentralized web app development that are based on Linked Data principles. The goal is not to create a new WWW – it uses existing WC3 standards and protocols (HTTP, REST, HTML, etc.). Today, URLs point us toward readable documents, but Solid wants them to point to data as well. By using http addresses to situate each datapoint with the addition of three identifiers (object: link source, predicate: link type, subject: link target) a sea of relationships between the data we create and make available can swell. The Web would no longer only offer us instances of answers to questions – it would be possible to compile our own, thanks to the “protein” (data) we feed it. With the support of Inrupt, a startup founded by Berners-Lee, users will be able to create their own personal online data stores (PODs) to house and maintain control over their data.

TORONTO: FROM SMART CITY TO TRUSTED CITY?

The City of Toronto, in partnership with Google-affiliated SideWalkLab, recently launched a project called *Smart City* in its Quayside waterfront district. There are a range of topics on the project’s agenda, including mobility, public space, housing, sustainable development, and digital innovation. This last theme has five key objectives: developing coherent and accessible digital infrastructures, creating new data standards, identifying responsible uses for data, and launching a core set of digital services that benefit the city

The Smart City project has, however, elicited some criticism. First of all, despite highlighting its desire for inclusiveness, the city’s inhabitants feel that they have been excluded due to a lack of information and little communication. Concerns have also been raised about personal data storage and use: Which data will the city be using... and for what? It would be necessary for the city to somehow differentiate between personal data and the anonymized urban data the city is likely to use. To clarify the issue, SideWalkLab designed visual language – deemed highly insufficient by some – and posted signs to explain the data capture devices dotted around the neighborhood so that citizens could see what they are used for, which data they har-

vest, which data will be identified, and whether individuals can be individually associated with those data.

In 2018, SideWalkLab suggested establishing a [Civic Data Trust](#) — a data management system that provides independent third-party data management, including collection, maintenance, anonymization, and sharing oversight. To complement this, Responsible Data Impact Assessments (RDIA, in-depth review and analysis) will be triggered whenever a proposal for the collection or use of urban data is presented, including details on the proposal's objective, its required data sources, any potential impact on individuals or communities, plus a risk-benefit analysis.

Mistrust of Google has led residents and city actors to question the role of the giant in the project's data governance. Would data really be “shared” with Google at the helm? The Toronto Area Chamber of Commerce doubted it, and [suggested](#) that the role of Civic Data Trust manager be given to the city library, because its brand is trusted, it is “neutral,” and it has expertise with archiving, data management policy, and information management issues.

After launching the initiative and sparking the resulting debate on data management, SideWalkLab's legitimacy — it is backed by Google, the epitome of ethically-questionable data use after all

— was put into question. Are they really suited to the role of mediator in the debate, and further, can a project of this size really be trustworthy? Strictly speaking, the Toronto project is not really an ideal example of Self Data at work, but it is extremely instructive in terms of the potential partnerships a city can put in place, and the role(s) that can and should be taken by citizens in the process.

LONDON: CREATING NEWER, MORE TRUSTWORTHY COLLECTIVE STRUCTURES

The City of London is also a member of the [Cities Coalition for Digital Rights](#), and has been exploring municipal data use for several years now. In 2010, it launched the Datastore, which stores data and makes it easily accessible for anyone to use in a variety of ways. In 2016, the city joined the Sharing Cities program, which promotes the development of Smart Cities and citizen engagement in their development. Along with the Open Data Institute, they are working towards clarifying what a data trust is and how it might be integrated into the Sharing Cities project — making it the terrain for future data trust exploration.

The London-based Overseas Development Institute (ODI) was co-founded in 2012 by Tim Ber-

ners-Lee and artificial intelligence expert Nigel Shadbolt, to showcase the value of open data and advocate for innovative ways of using data to inspire positive change at the global level. It is a non-profit, non-partisan, independent organization. A data trust could feasibly be presented as a safety measure, and their use a sign that data transfers between holders and organizations requiring them will be handled securely. Using them would create new economic opportunities, help with research, and greatly empower communities. The definition given by the ODI, however, is limited to “a legal structure that allows independent management of data by a trusted third party.”

The ODI relies heavily on the research of [Sean McDonald](#), who defines a data trust ecosystem as having five dimensions: a creator/owner, a beneficiary, a trusted third party, an asset, and a goal. An owner gives an asset to a trusted third party to ensure that it fulfills a useful function for the beneficiary. However, according to McDonald, data trusts are not a panacea — they carry their own set of risks, including their potential to foster tax fraud, the possibility that actors may divert data to serve commercial ends or even waive responsibility for data privacy. The data trust governance model cannot do without a political and legislative system supporting it.

The ODI has set up three data trust implementa-

tion pilots since 2016. The first, in partnership with WRAP UK (Waste & Resources Action Program), tracked the impact of a data trust model on food waste by improving stakeholders' ability to monitor and to measure food waste in supply chains.

The second pilot was conducted in partnership with WILDLABS Tech Hub. The collaborators spent the first three months of 2019 exploring whether a data trust model could be used to scaffold the efforts of the numerous developers, vendors, users, and data re-users working to tackle the illegal wildlife trade around the world.

The third pilot explored two use cases, each applying the data trust model explicitly in the context of the city. One dealt with mobility, and how using a data trust could maximize available parking data and promote more attractive, less polluting means of transport. The other looked at ways to improve energy efficiency in one of the city's blocks of social housing flats, by installing sensors to monitor and control a modernized collective heating system.

Sharing Cities

The Sharing Cities program began in 2016. It brings together 100 European communities, including London, Bordeaux, Lisbon, Burgas, Warsaw, and Milan, to create a shared testing ground and a shared methodology for smart cities design that integrates citizens into the process.

One of Sharing Cities' objectives is to create Urban Sharing Platforms, which will enable users to manage data from a wide range of sources (sensors, analytics, etc.) using a set of common principles and open technologies. It leverages London's expertise in data analytics (DataStore), Milan's work in with application programming interfaces (APIs) and with public data use, and Lisbon's experience analyzing sensor data and data gateways to achieve this goal.

TRENTO: USING SMARTPHONES AS DEMOCRATIC TOOLS

Telecom Italia's *Mobile Territorial Lab* began a Self Data initiative in 2012, in collaboration with Michele Vescovi, a researcher at SKIL Laboratory (Semantic and Knowledge Innovation); the Bruno Kessler Foundation; and Telefonica I+D. Efforts have been centralized around the creation of [My Data Store](#), a tool enabling users to control and share their personal data. The Mobile Territorial Lab (MTL) aims to create a “living” lab at the heart of real world Trento, Italy. One of its specific aims is to harness mobile phone detection capabilities to track and make sense of families' spending and lifestyle behaviors, plus markers indicating mood and stress patterns.

My Data Store is a tool for the control and sharing of personal data collected from mobile phones and via experience sampling apps. The project was tested with 63 participants over 15 weeks; each received an Android smartphone equipped with mini-detection software that aggregates data detected from the mobile phone on the user's My Data Store account. The data is securely stored for future exploitation.

The data come from three sources. Firstly, there are those automatically extracted from the mobile devices' call logs and SMS messages and

from scans of Bluetooth devices and GPS/WiFi, multimedia activity events, and device status information. Secondly, the devices are equipped with sensors that periodically detect air quality and basic meteorological data (temperature, humidity, etc.), which is then transmitted to the Store. Finally, some data on daily mood, stress, sleep quality, and daily expenses are collected using an experiential survey conducted via mobile app. The findings revealed that close to 50% of participants said that they gained a greater awareness of what personal data is, how they are used, and what their potential might be.

Other sources of personal data were shared on the MyData Store platform, and many dedicated apps were developed that enabled citizens to reuse them. For example, a strategic partnership with the Coop distribution chain enabled Trento's residents to retrieve their consumption data and reuse it (in combination with their mobility data) on third-party apps designed to manage their shopping routines.

HELSINKI: THE CRADLE OF MYDATA

Finland (especially Helsinki) has been the cradle of Self Data – known synonymously as MyData – since 2010, when the *MyData Global* network was created. Fing is a founding member. On July 1, 2019, when the country took over the presidency of the European Union, their agenda makes mention of their intention to accelerate the development of a human-centered data economy during the 6 months of their presidency.

We are not featuring Helsinki here to highlight another city's approach to personal data. Rather, we seek to draw your attention to what is now an indispensable tool for any city wishing to embark on Self Data implementation: membership in the MyData Global network (whose epicenter is in Helsinki). Its members are making a concerted effort to recalibrate the personal data economy more equitably in favor of individuals. In October 2018, the network was officially formalized as a nonprofit.

MyData Global brings together a wide range of entrepreneurs, activists, academics, private companies, and public organizations. Its aim is to centralize emerging reflections and increase their impact at the local level, and in doing so, enrich perspectives, capabilities, and experience. The network seeks to “empower individuals by improving their right to self-determination regarding

their personal data,” especially by encouraging others to launch pilot research programs. The community meets annually in Helsinki for the MyData conference.

The Finnish headquarters is charged with organizing the annual conference in Helsinki, to create a space for discussion and increase the visibility of local hubs' concrete activities. Each hub is built according to the MyData Declaration; their mission is to form a local community, define a common goal, and work collectively on project implementation. Fing is responsible for the MyData France hub and the French community working on the topic.

Local hubs are growing rapidly in the cities of Atlanta, Barcelona, Brussels, Geneva, London, Zurich, and Sydney; and nationally in Austria, Brazil, Cameroon, Denmark, Greece, Hungary, Japan, the Netherlands, Scotland, the United Kingdom, Sweden, Slovenia, and Silicon Valley.

The network is a precious resource for local public actors for a number of reasons. MyData's sphere of influence is gradually expanding (particularly at the European level) by capitalizing on the efforts of existing networks such as Eurocities', whose Citizen Data theme is a driver of its activities. Published in March 2019, iits Citizen Data Principles state that “governments have the responsibility and must ensure that citizens can have access to

and manage their data [ie: MyData] as well as influence how it is collected and used.”

Via the “Virium Helsinki Forum,” Helsinki has begun to reflect on ways to effectively implement the MyData/Self Data model. The city has already begun mapping the sources of personal data it holds and developing concepts for ways to enable individuals to retrieve and reuse them. The data collection report shows that at least 209 of the approximately 800 Helsinki city computer systems contain personal data.

The city is currently working to become a MyData operator. It is developing demonstrations to illustrate how it might be possible to manage citizens' energy consumption via connected meter data; while its project CaPe, funded by EIT Digital, is working towards making mobility data portable so that third parties can create services that are useful to individuals.

THE CITY OF NANTES, GREATER LYON, AND LA ROCHELLE: FRENCH CITIES AT THE FOREFRONT OF SELF DATA

Frances' efforts to implement Self Data, spearheaded by Fing, have spread separately to several cities over the past few years. We have been working directly with three French cities – Nantes Métropole, La Rochelle and Greater Lyon – on the design and implementation of their Self Data Cities experiments. The results are presented in the following chapter. The cities of Rennes and Grenoble are also engaged in data sharing initiatives, which we present here.

Grenoble

Grenoble-Alpes Métropole is in the process of reconsidering how it manages and shares personal data. During a conference at Fing in March of 2019, the city's digital and Smart City project manager Laurent Deslattes outlined the core areas of the city's internal investigation: which data do we hold? How are we using them, and how can we use them? Where do we place citizens in this process? What are the provisions that will optimize data use by and for citizens, and respect their privacy?

Out in the field in Grenoble, La Turbine.coop – formerly known as Scop La Péniche – is a space where people can co-create collaborative tools, a digital data mediation Infolab, and a social innovation project accelerator. It is a place where citizens and users can take charge of their digital and data challenges and collaboratively invent digital solutions that contribute to transitions. In 2015, La Turbine.coop held and co-hosted Fing MesInfos Energy workshops along with its participation in several projects exploring individual control over personal data, including the MétroEnergies (formerly Vivacité) project. The MétroEnergies platform, now managed by the City of Grenoble, generates tailored visualizations depicting individual energy data (energy consumption and associated expenses), challenges users to lower their energy consumption, generates comparative analyses with other users data, and facilitates data export. Its activities will eventually augment the contents of energy maps needed for territorial energy planning, help to define and compare baseline energy profiles, serve as a database for research, help to identify and address energy poverty, etc.

Captothèque is another initiative being led by the city of Grenoble. Its citizen sensing and participation scheme gives inhabitants a set of tools (micro-sensors, applications, smartphone, dedicated website, etc.) so they can collect air

quality data that will enrich the city's existing indicators. Their Citizen Data Science approach is similar to those developed by Barcelona and Amsterdam, for example. Today, the city is considering integrating Self Data logic into its regional data warehouse strategy, so that individuals can control and share their personal data themselves.

Rennes, France

On a different tack, the Academy of Rennes has been focused on implementing Self Data principles in the education sector. The reasoning is this: students' and teachers' paths are fragmented through the use of multiple educational platforms (eg: software such as pronote, Moodle, Pearltrees, PMB, and POD; news reports shown in class; digital textbooks, encyclopedias and dictionaries; France's digital resource bank BRNE; the national education portal eduthèque; the national extra-curricular education pathway platform Folios; France's online curriculum platform Labomep; homework websites like Homework done; pedagogical platforms such as Toutapad; digital magazine creator Madmagz; Satchel teaching and learning tools; online digital certification program PIX; mapmaking platform Cartoun; VIA, AALN, M@gistère, etc.) and workspaces (school, libraries, home, etc.), not to mention when they change schools – which renders a full and complete academic record virtually impossible to generate.

With its MyToutatice project, an extension of ENT Toutatice, the academy intends to make students' school records much easier to read. MyToutatice is personal cloud platform designed in collaboration with CozyCloud – a “digital binder” that can collect and store learners' and teachers' data that will endure until the end of their schooling or teaching career, even if there are pauses along the way. It allows users to control and exploit their personal data, while being respectful of their privacy. An experiment involving high school students and teachers is expected soon.

The City of Nantes, La Rochelle, and Greater Lyon

Three French cities have become pioneers of Self Data. They have been working closely with Fing and its partners to achieve an ambitious objective: launch experiments that to develop their understanding of what Self Data implementation implies for their regions. La Rochelle, the City of Nantes, and Greater Lyon are launching Self Data experiments focusing on sustainable mobility, the energy transition and social welfare, respectively. This is the first time that local public actors have positioned themselves as leaders of the Self Data revolution! The next chapter details their actions.

Project RUDI

The Rennes Urban Data Interface (RUDI) project, initiated with the support of the European Union and the city of Rennes (in conjunction with its Metropolitan Public Data Service), will launch in September of 2019. It will develop an urban data interface to facilitate access to and understanding of the data it contains, principally to stimulate innovation in the region. Operating as a kind of “social data network” for citizens, social enterprises, and businesses, it will operate as a development and proving ground for new services that take advantage of access to these data and to data furnished voluntarily by the city's inhabitants. In the true spirit of Self Data, citizens are at the heart of the undertaking – they should eventually be able to access personalized and also anonymized data flows, manage their consent to share their data with third parties (services or entities), and gain access to services that reuse their data.

03

**SELF DATA DIY – CREATE
A LOCALIZED, RELEVANT
SELF DATA INITIATIVE
FOR YOUR REGION: THE
EXAMPLES OF THE CITY
OF NANTES, LA ROCHELLE
AND GREATER LYON.**

03

SELF DATA DIY – CREATE A LOCALIZED, RELEVANT SELF DATA INITIATIVE FOR YOUR REGION: THE EXAMPLES OF THE CITY OF NANTES, LA ROCHELLE AND GREATER LYON.

What about your city joining other Self Data Cities? Before embarking on a Self Data experiment, following a few preliminary steps will help you not only create the necessary dynamic in your region – you will also be able to more clearly foresee some of what you will encounter during implementation. We went through each of these phases with all three cities; the results and methodologies this kit contains are intended to help you do the same.

Between September 2018 and July 2019, we worked with three cities to help them :

- » **phase 0 : become aware of Self Data and choose a challenge that would be appropriate to their city's needs;**
- » **phase 1 : define the perimeter of the personal data that could be shared;**
- » **phase 2 : create use cases and their associated governance models;**
- » **phase 3 : finalize a road map to implement one or more experimental scenarios in the coming year.**

UNDERSTANDING THE STAKES ASSOCIATED WITH SELF DATA, AND CHOOSING AN APPROPRIATE PROJECT

The first step to launch Self Data is to rally your accomplices for the coming year: colleagues, data holders, civil society representatives, business clusters, universities, etc. You will need them to help build an experimental scenario that a majority of them will adhere to.

You will not be able to cover every area, nor will you manage to unite actors from every single sector, so you will need to find a topic and theme that sticks. Project your minds into the future: which themes are the most likely to motivate internal and external actors? Which of your municipal activities would most benefit from taking a Self Data approach? In La Rochelle, for example, the city settled on sustainable mobility as their theme for several reasons. First, the city is already a mobility pioneer. Second, its energy and environment agency ADEME was keen to pursue mobility. Then, a meeting with the mobility accelerator Fabrique des Mobilités, who were considering creating a “mobility account” based on Self Data, prompted a joining of forces. And finally, La Rochelle wanted

to set up a regional “carbon calculator” that could incorporate residents’ carbon footprint data with their explicit consent and under their control.

Even though Greater Lyon initially gravitated towards education as its theme, given the amount of data the city holds and its potential uses, but they opted instead to explore social welfare, given the city’s focus on solidarity (French only). The topic tied together several other questions the city had around its administration and its citizens ability to grasp and exercise their rights. The city had, in fact, already begun to explore the terrain from a digital perspective, for example by using France Connect as the interface between parents and public childcare services.

As for the city of Nantes, the situation was a little different. The people we worked with were neither information systems department employees nor did they work for the city’s innovation department – they came from its business divisions, including the DGTEESU (Directorate General for Energy Transition, Environment, and Urban Services). The social welfare theme was a great fit, as were its focal areas: energy and food.

Got your theme? Then it’s time to identify a core group of local actors who either know the field,

or are working with data – especially personal data. These key associates will support you in the coming months as you map out your Self Data strategy. Many of them will not know Self Data; it is a complex subject. You will have to organize an orientation day, invite them, and then use the opportunity to further sharpen your area of focus. That Self Data seminar must achieve several objectives:

- » participants must come away with a solid understanding of Self Data and its potential;
- » participants must be able to tie Self Data together with their specific local concerns;
- » participants will have the desire to pursue the topic of Self Data further, and consider data and data uses;
- » you have a list of specific challenges for Self Data to respond to;
- » you have drawn up the “guest list” for future workshops, when you can open the debate to a wider audience, conduct interviews, have meetings....

You will find a methodology and a set of best practices to help you prepare your seminar in **appendix 1** of this booklet. You may also wish to

consult the [Self Data FAQ](#), so you are prepared to answer your participants’ questions. There is also the [Self Data Pilot Summary](#) (especially Chapter 4, which is full of examples) and [all the other](#) (English) content we have produced over the past 7 years. Everything is available under a Creative Commons 3.0 attribution license, enabling you to reuse the info as you wish, as long as you cite us as the source.

In Nantes Métropole, Grand Lyon and La Rochelle, each of the seminars we conducted helped us to get a clearer picture of the challenges to be overcome at every stage in the project process.

Nantes Métropole - Self Data in support of the energy transition, and to help me...

- » **...understand and take charge of my food choices and their impacts;**
- » **...contribute to clean energy production in my neighborhood or city;**
- » **... reduce my home's carbon footprint and manage my energy consumption.**

Greater Lyon - Self Data to improve social welfare, and to help me...

» ...understand my entitlements, streamline my tasks and responsibilities;

» ...easily take part in the city’s public sports, cultural, and social activities;

» ...identify the right people to talk to, and perform necessary tasks.

La Rochelle - Self Data for better, more sustainable mobility, and to help me...

» ...calculate and reduce my mobility carbon footprint;

» ...manage my mobility budget;

» ...contribute to rethinking the mobility options on offer in the region;

»combine utility with pleasure: reduce the carbon footprint of my daily comings and goings, and enrich my cultural knowledge.

DEFINING THE DATA PERIMETER

Step two of the Self Data adventure: identify the relevant personal data that will eventually be shared with individuals. This happens during what we call a “datablitz” or a “data hunting workshop.” Based on the challenges you have decided to address, three subsets of questions must be answered:

» **Which are the most relevant data for this topic? (maintain your focus on personal data, but also look at data repositories and open data which might be useful).**

» **Who holds these data? Do they already facilitate access (via API; direct download)?**

» **What are some useful ways to reuse or combine these data?**

Appendix 2 details the datablitz methodology. These workshops are extremely useful opportunities to explore and systematically identify sources for personal

data that perhaps did not come immediately to mind, especially data from other sectors and under different topic headings than your own that still make sense within the context of Self Data implementation.

From your workshop findings, you can create a shared spreadsheet containing the results so that anyone – but especially the workshop participants – can contribute to it. We have created an online [spreadsheet mapping tool](#) (French language) that generates “data sheet” visualizations for educational purposes. It can be further enriched after a questionnaire is completed.

You can filter your search by topic, data type (personal/non-personal), or keyword. Each data sheet presents information title, the data, the holder(s) and the platform. There are over 60 types of mobility data, and 100 types of data relevant to the



energy transition. Two-thirds of these are personal in character. These include profile data (age, name, address, household composition, etc.), geolocation data, consumption data (subscriptions, fuel, food, energy), preferences data (specific itineraries, food, etc.), and household characteristics. These are the type of data that are equally useful when determining which solar panel subsidies someone is entitled to, or to enable people to make a list of things to buy from sellers who offer local, seasonal products that suit their needs and preferences (budget, tastes, etc.).

The aim of a datablitz is not to come up with an exhaustive list of the personal data you wish to share with individuals. Once underway, part of your experimentation process will be to carefully document the data in each holder’s information system (IS) anyway. These workshops are an opportunity to expand your horizons about what could or might be shared with people from the perspective of existing challenges – not the information systems at your disposal. You might be surprised to discover new types of data and new data holders that were not on your radar previously.

POTENTIAL USE CASES AND RELEVANT MODELS OF GOVERNANCE

Gathering and visualizing the data on a data-sheet also puts you in a position where you can begin to imagine new uses, services, and use cases for the data.

The third stage of the project process is to dream up those new use cases and their relevant governance models (personal cloud, civic trust, citizen’s right to portability, metropolitan data platform, etc.). It’s a lot of work – and you will only have a couple of workshops’ worth of time to dedicate to the task. During the first you will come up with at least a dozen service concepts, and then choose a few based on the datasheets, which you then flesh out into use cases. This is not the time to let your imagination be constrained by technology, architecture, or data sharing concerns. All of those will be dealt with during the second workshop, during which you will develop an overall approach to implementation and governance for each use case. You will need to familiarize yourself with the “on-the-shelf” options that exist (see Chapter 1 of this kit), create an alternative, hybridized model that can support a use case that

itself overcomes at least one of your inhabitants’ challenges. Appendices 3 and 4 describe the methodology you can use to organize and host these workshops.

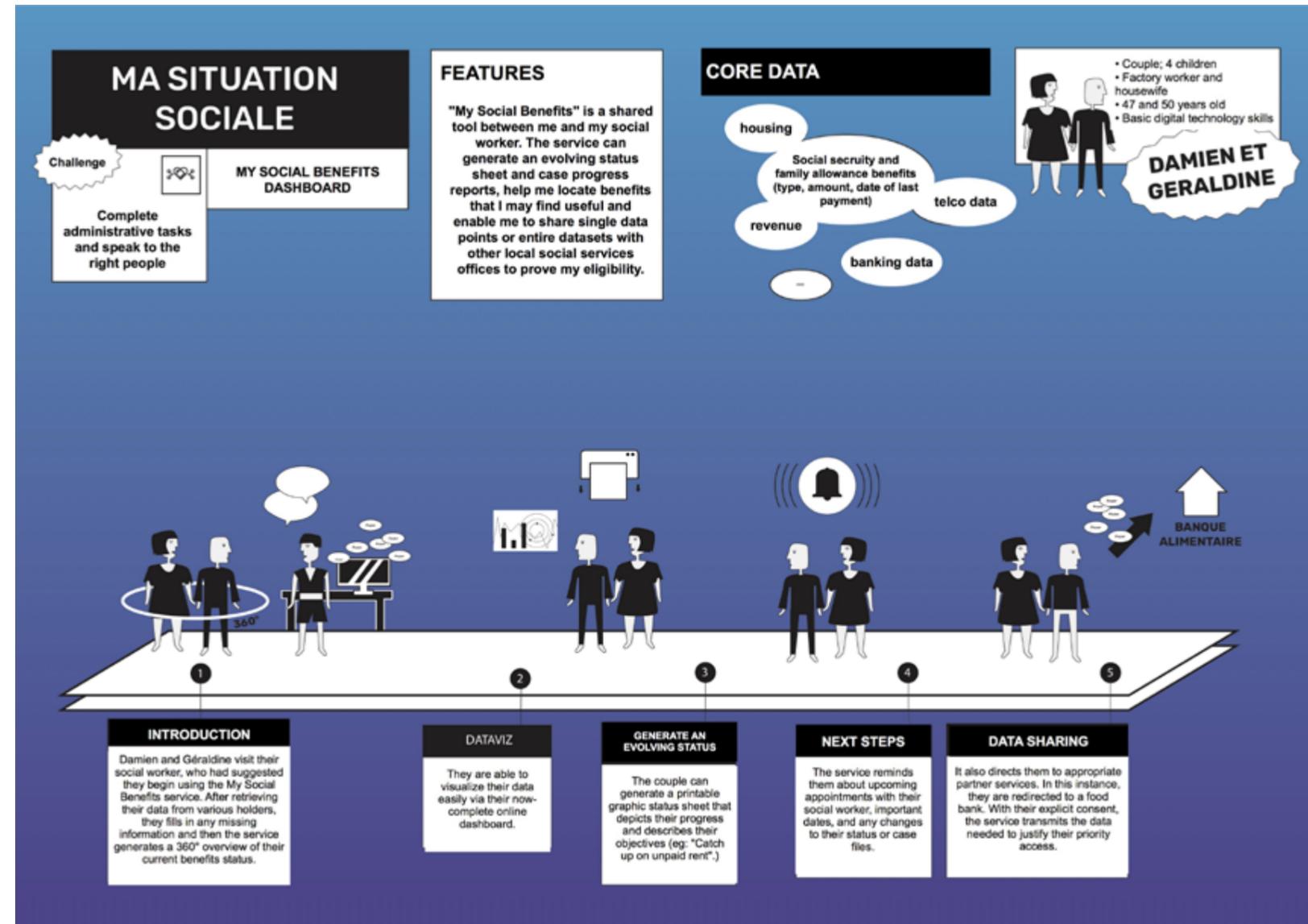
Our work with Nantes Métropole, Greater Lyon, and La Rochelle yielded a dozen use cases – each of which includes details about the data utilized, plus their features and target audiences – as well as a use scenario and the governance models that might be used to support them. In the following pages, we present a brief synopsis of each.

1

MA SITUATION SOCIALE - MY SOCIAL BENEFITS DASHBOARD

SELF DATA AND SOCIAL WELFARE - GREATER LYON

CONCEPT AND USE CASE

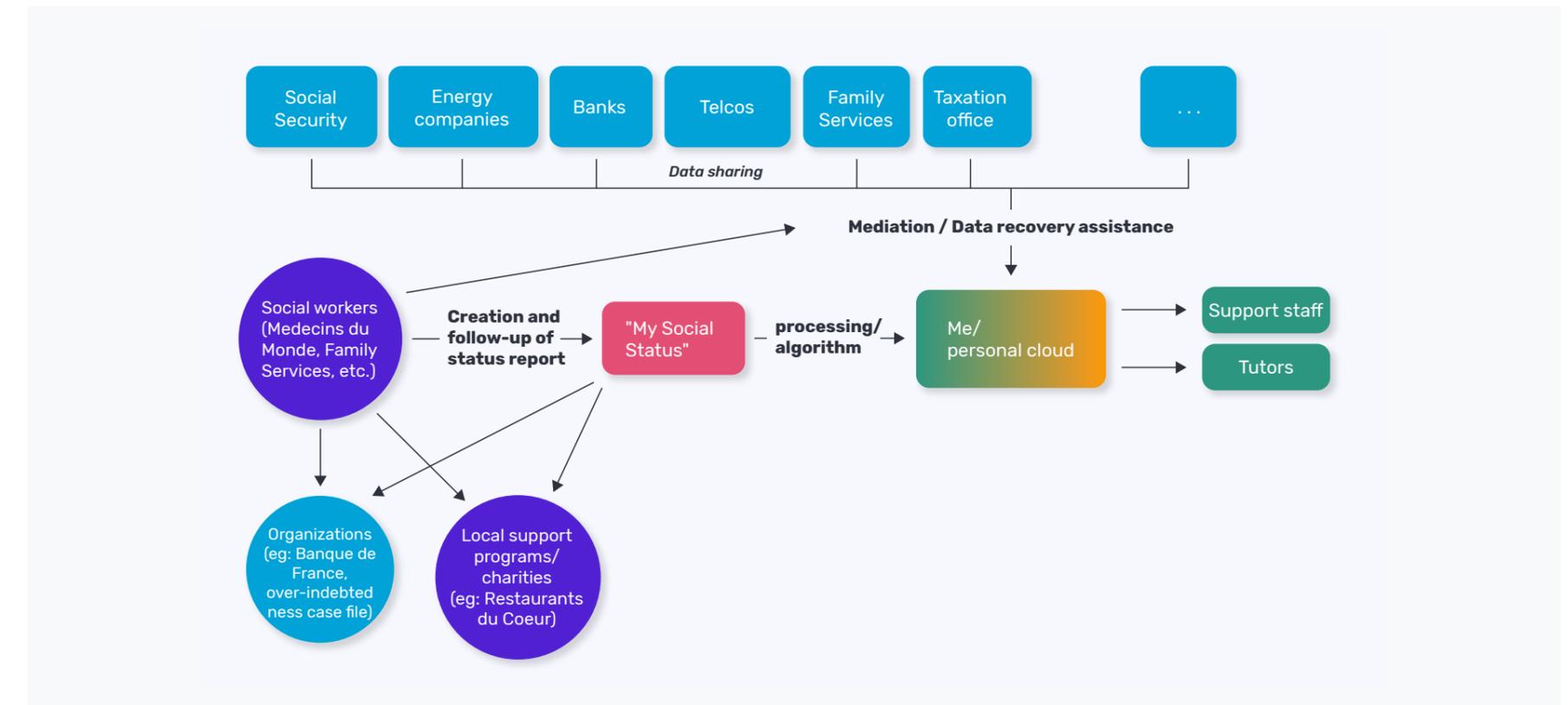


GOVERNANCE MODEL: THE PERSONAL CLOUD

This tool is destined largely for social services beneficiaries who are already seeing a social worker. "My Social Benefits Dashboard" aligns well with the personal cloud configuration, as the personal cloud supports multiple data source interoperability and individual storage in a digital "home" where users are able to store and administer their personal data. This is a necessary prerequisite to its operation, but

this third party service is actually intended to enable data visualization, so users can assess their benefits status. The visualization is shared, built for individuals and for the social workers who play a major role in gathering data for the evaluation, and then using it to generate the status summary sheet, case progress report, and a list of the beneficiary's tasks and priorities. The service cannot operate if the "per-

sonal cloud" and "third party service" functions are siloed. The sharing of data with social services is required as is securing that exchange. Users' trust in the service's data security measures is capital, and for this reason, this type of service would best be developed by a public actor that already handles this type of exchange (the City administration, Family Services, etc.).

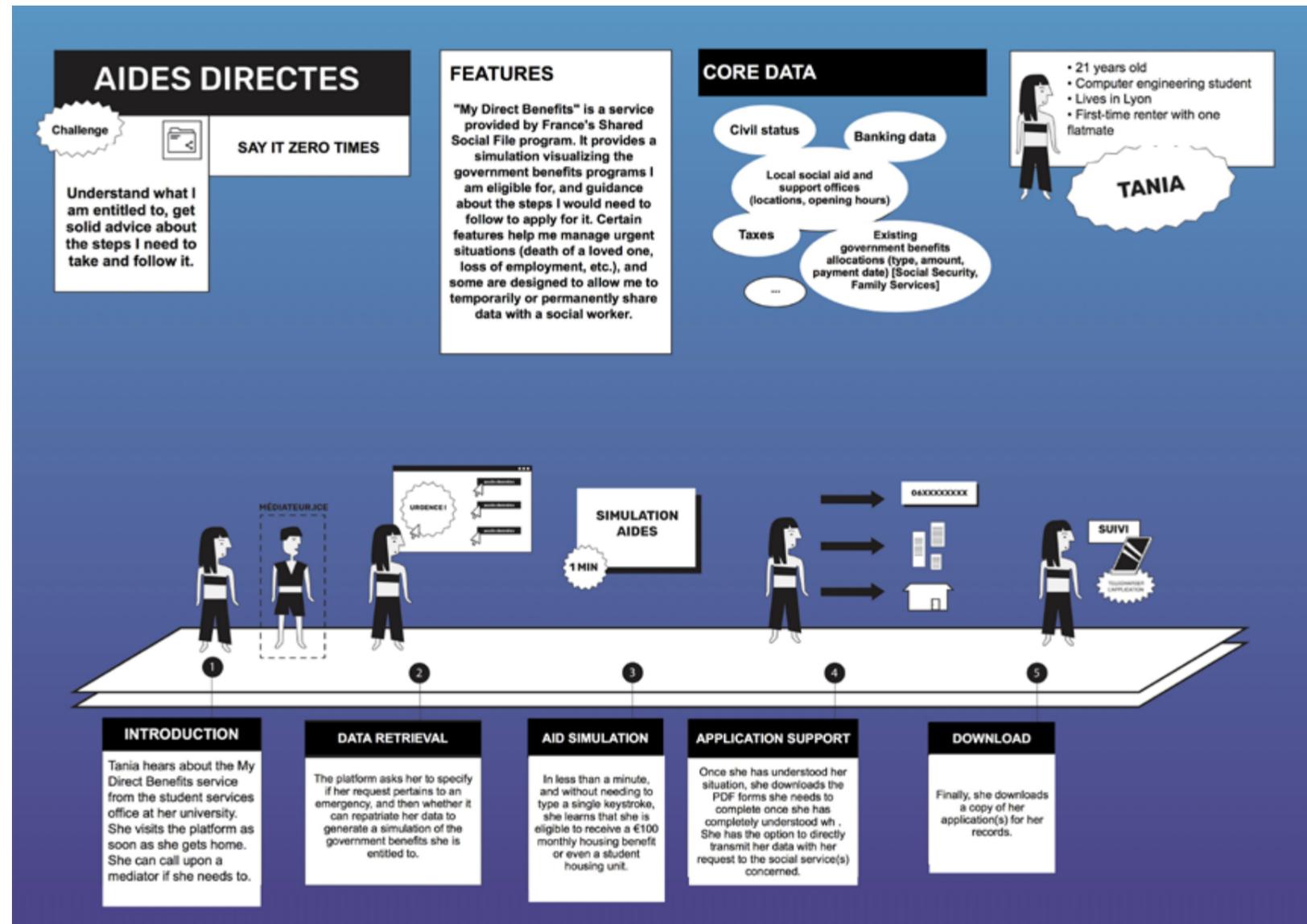


2

MES AIDES DIRECTES - MY DIRECT BENEFITS

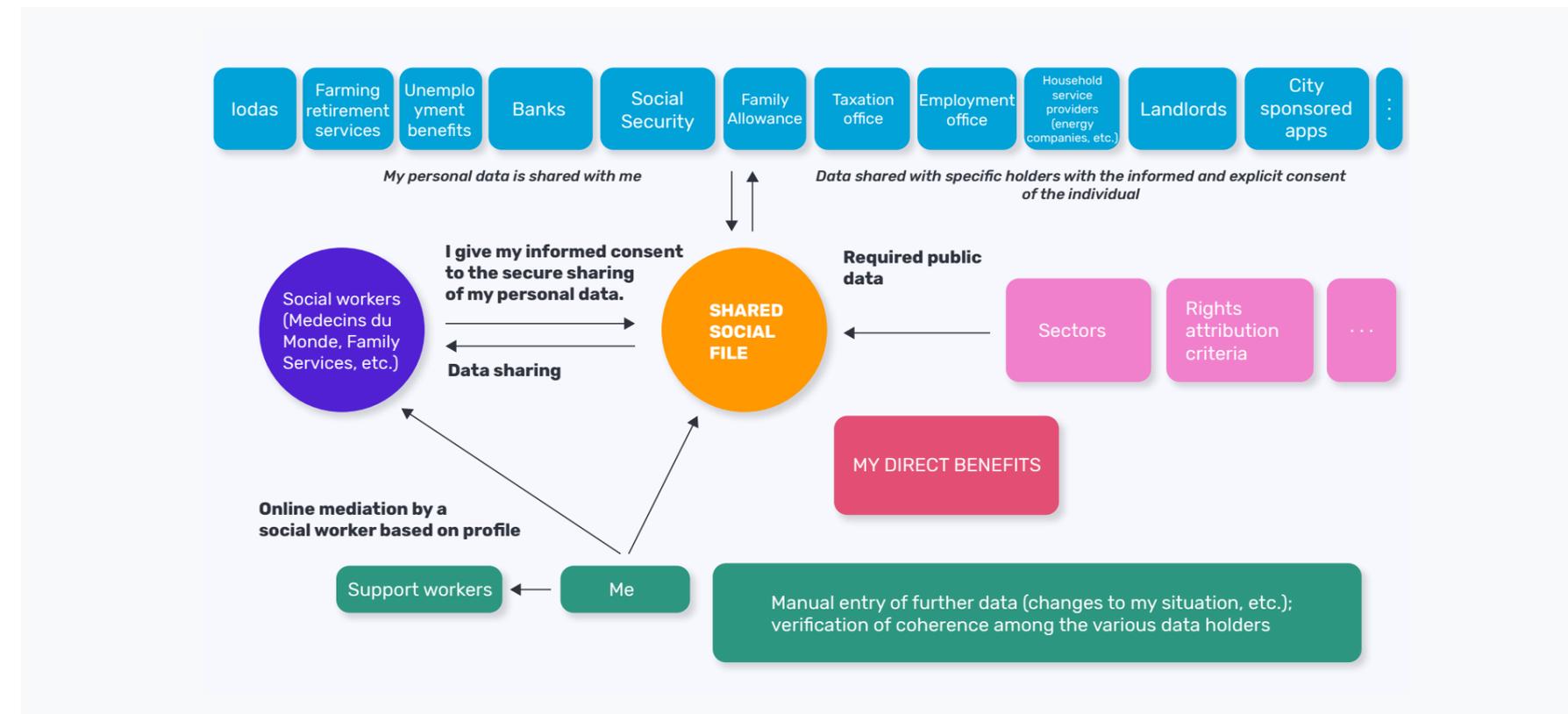
SELF DATA AND SOCIAL WELFARE - GREATER LYON

CONCEPT AND USE CASE



GOVERNANCE MODEL: THE TRUSTED THIRD PARTY PLATFORM

"My Direct Benefits" is an application that would be integrated into France's "Shared Social Folder" platform, which allows individuals to aggregate their data in a secure space and manage sharing. The service must be provided by a trusted third party (eg: Family Allowance services). The platform must be able to transmit data securely to social workers, whether temporarily (eg: with a temporary password) or permanently. In line with the Shared Medical Records model, the individual can only share data with other data holders (organizations, social workers, etc.) and cannot reuse them via third-party services. However, given the ways My Direct Benefits will be using individuals' data, third party reuse seems unnecessary. A purpose built service, bundled with the platform and provided by the same trusted third party organization, would likely be all users would need.

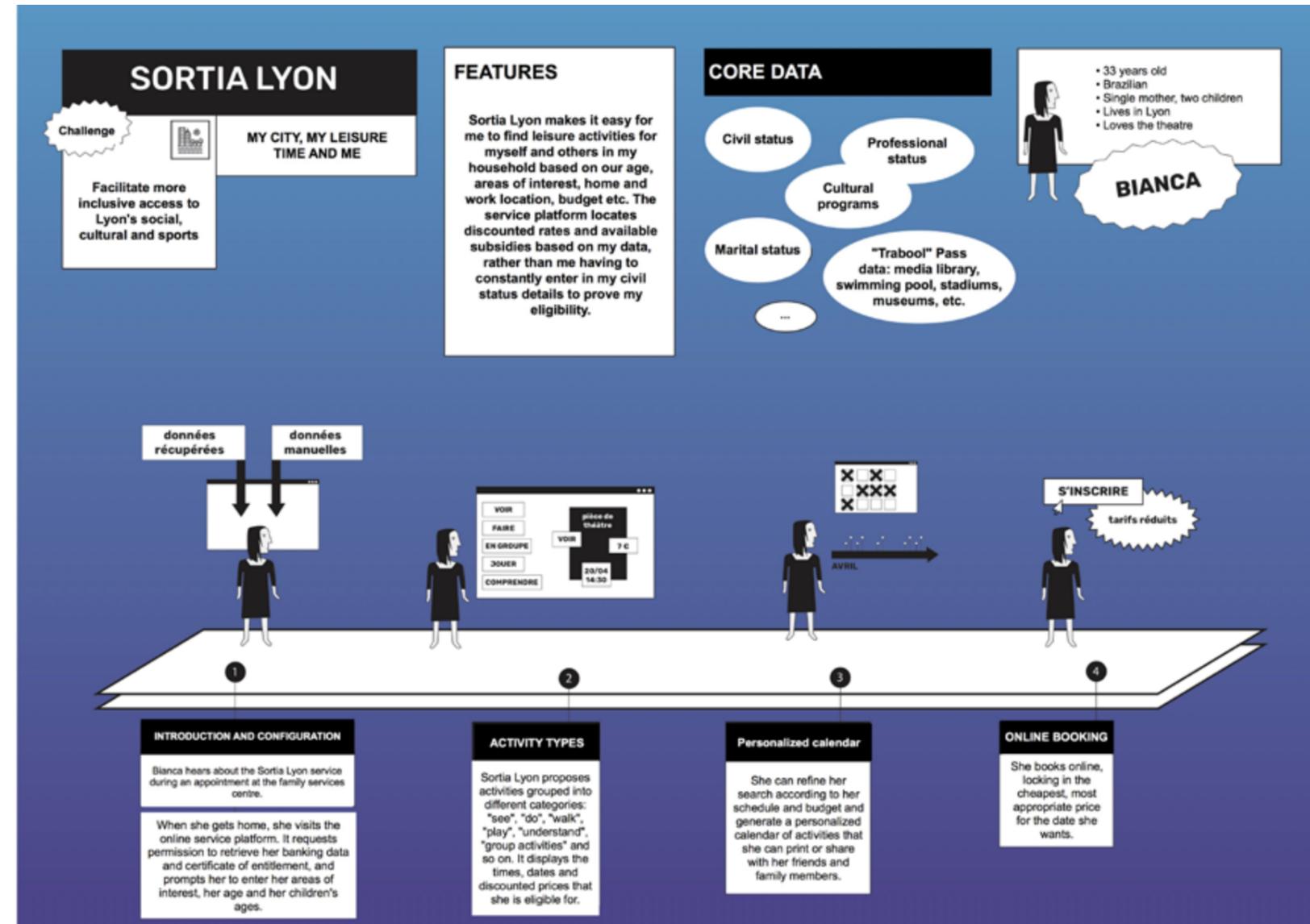


3

SORTIA LYON

SELF DATA AND SOCIAL WELFARE - GREATER LYON

CONCEPT AND USE CASE

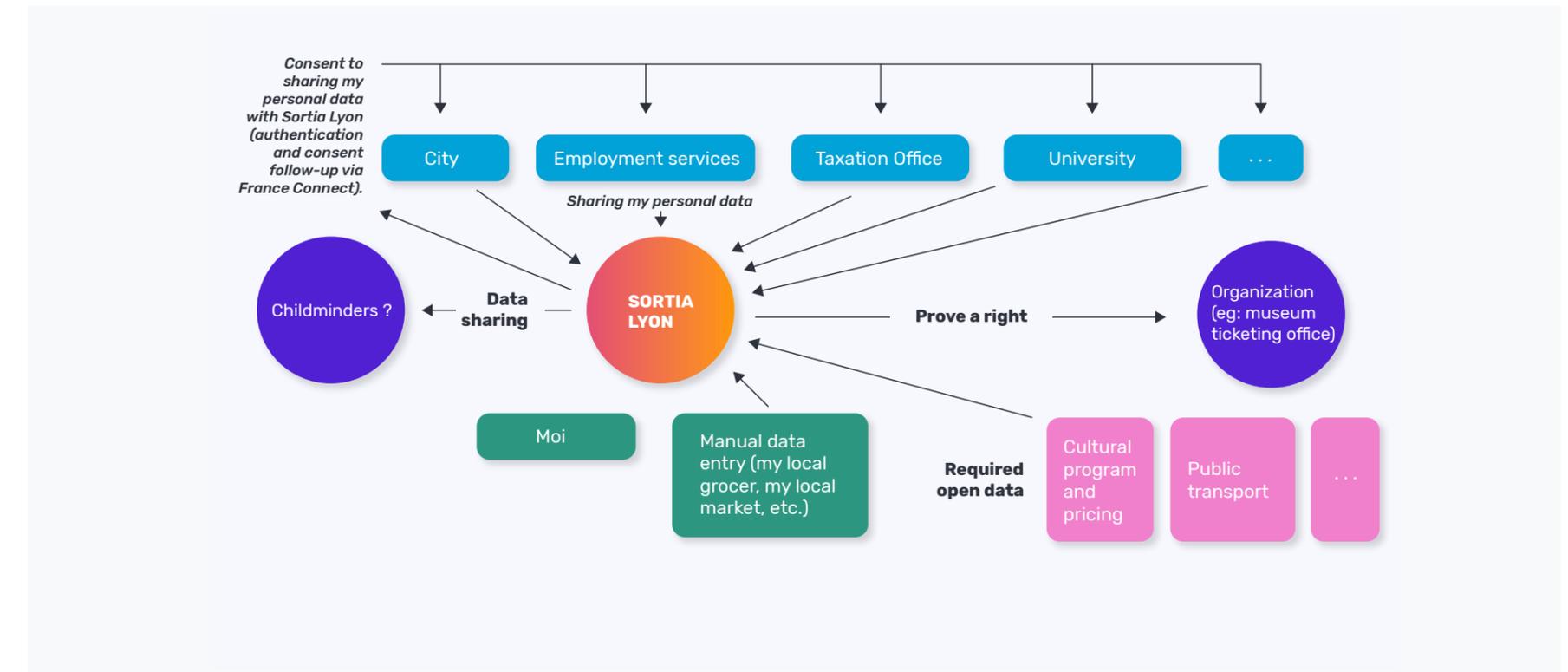


GOVERNANCE MODEL: DIRECT DATA TRANSFER

The direct transfer model applies if Sortia Lyon can operate with a restricted number of data sources. Given that most of their sources are public and already a part of France's national social services platform FranceConnect, Sortia Lyon could use those ties to streamline its user authentication and consent follow-up process. The "Direct Transfer" model does not really foster the creation of a truly perennial 360° data overview, so the choice

was made to share less external data with Sortia Lyon and ask the users to enter their preferences and interests. The service can operate in a vacuum and simply inform users of their rights, or it can also allow users to share some data with a ticketing agent, for example, to prove their eligibility for discounted rates without having to share the entirety of their confidential data. Creating a professional interface for childminders – so they may

facilitate families' access to cultural offerings and streamline preparations for outings, for example – is also a possibility.

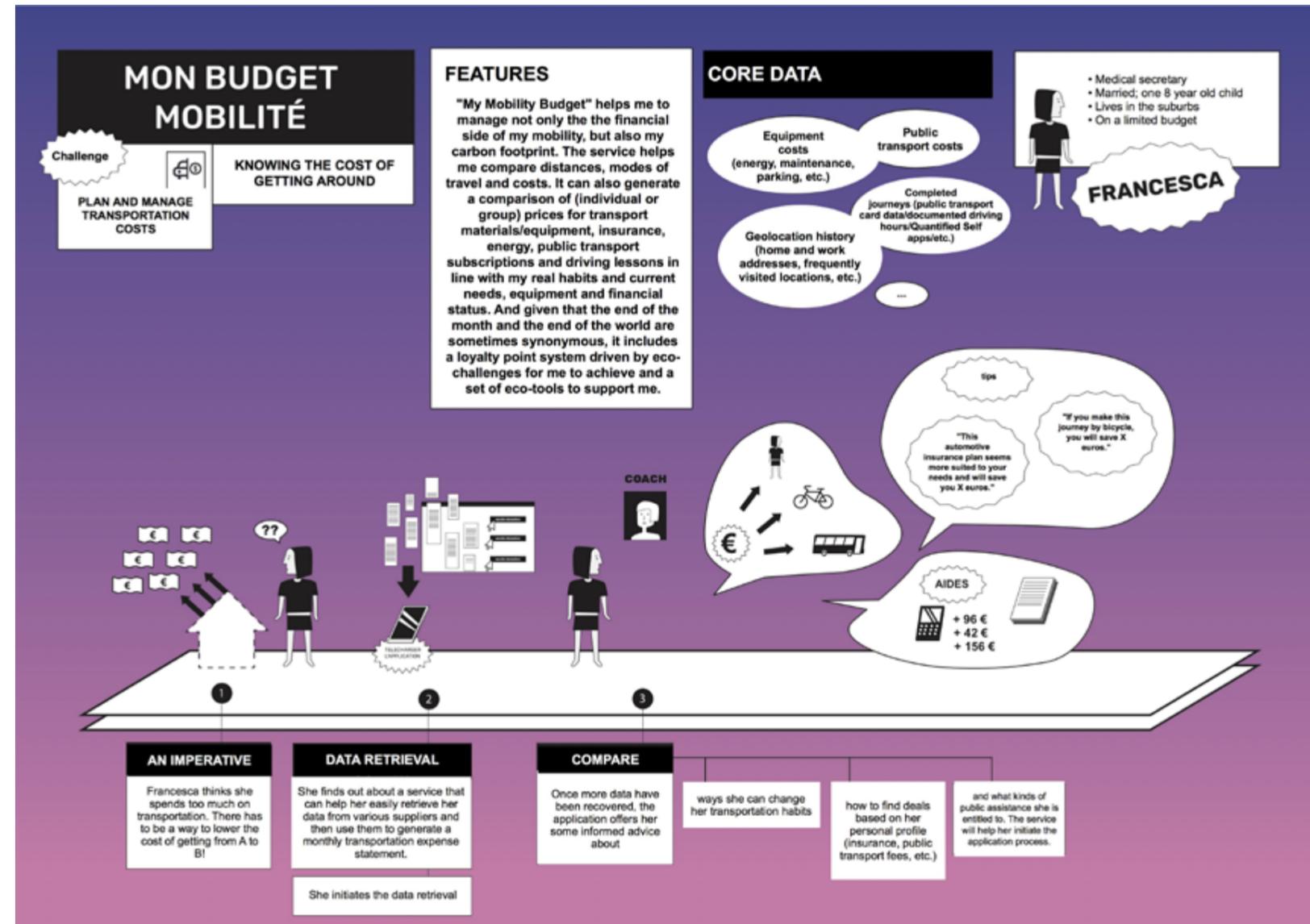


1

MON BUDGET MOBILITÉ - MY MOBILITY BUDGET

SELF DATA AND SUSTAINABLE MOBILITY - LA ROCHELLE

CONCEPT AND USE CASE

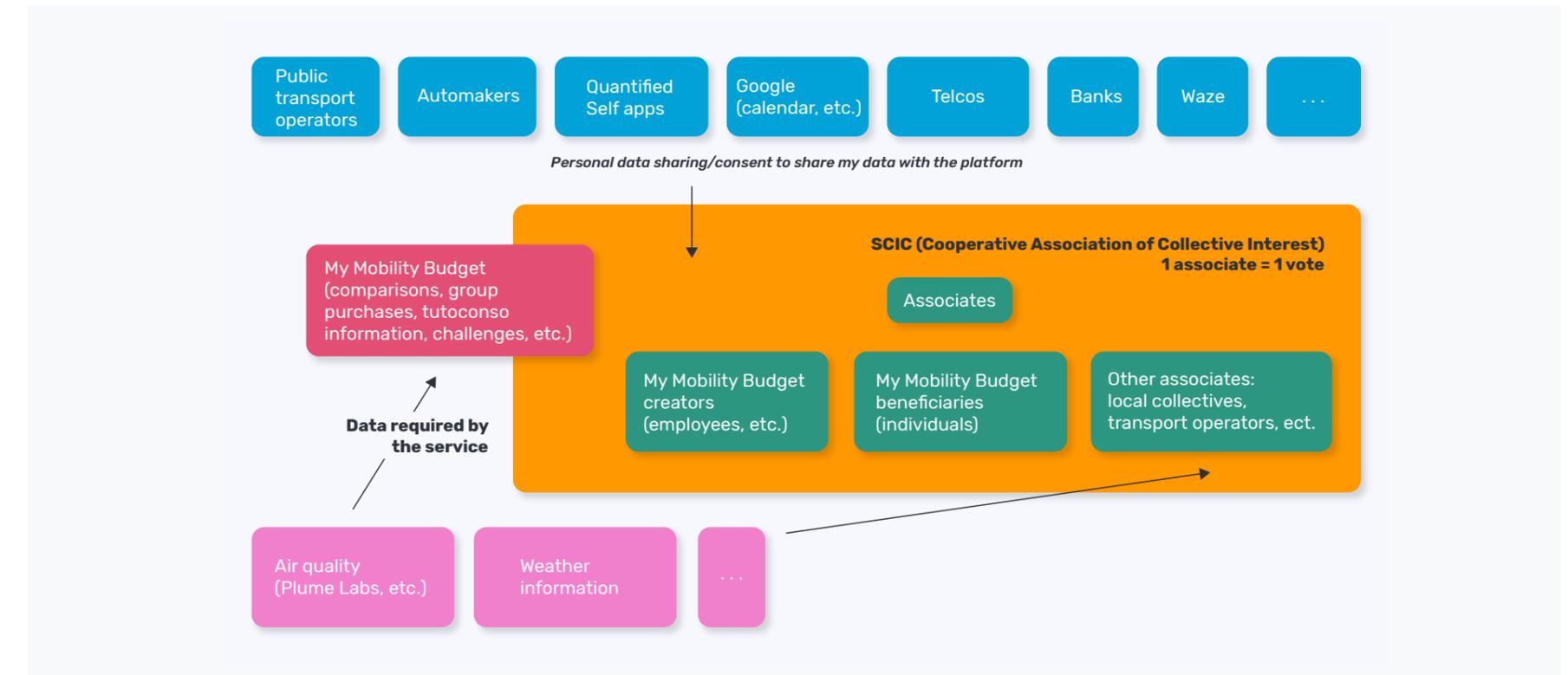


GOVERNANCE MODEL: DATA COOPERATIVE

The main problem with using a data cooperative to develop this kind of service is how difficult it can be to get people involved; there has to be a strong interest on the part of many individuals to commence building a structure like this, because of the time that must be devoted to governance. And achieving the critical mass among members required to produce a high quality digital service and preserve anonymity might be tricky. To overcome

this obstacle, the solution for My Mobility Budget could be to create an SCIC – a cooperative association of collective interest – and organize a collective. That way, there would be greater means with which to develop the service, critical mass could be more easily reached and the service developed could be used by any member of the collective, especially for things like making group (and thus discounted) purchases. The reasons for

a collective to join an SCIC are numerous: to harmonize public transportation offers, change citizens' habits in a positive way (via challenges, or a bonus system to encourage change – free bicycle storage, for example), or otherwise act in the interests of public health.

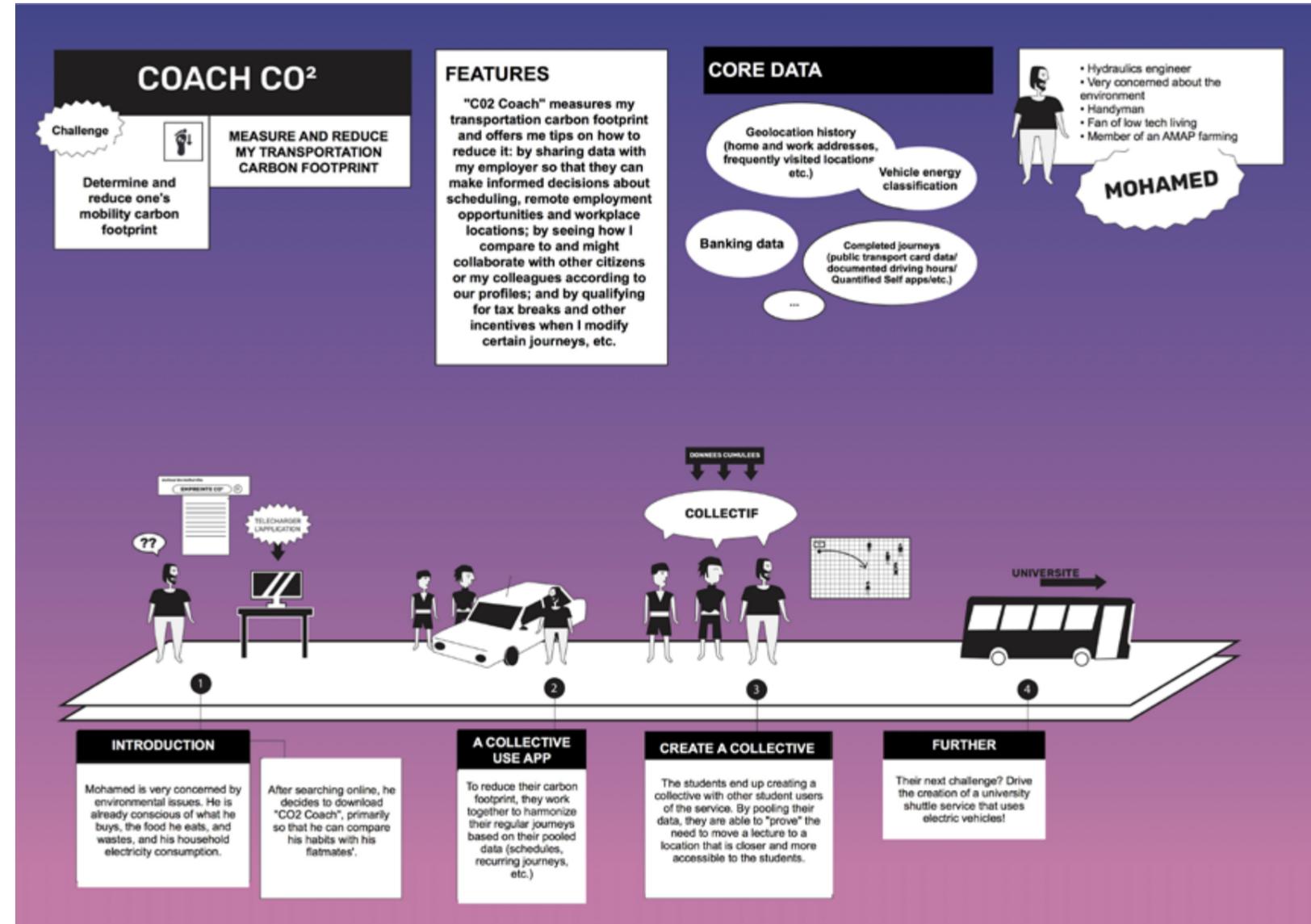


2

COACH CO2

SELF DATA AND SUSTAINABLE MOBILITY - LA ROCHELLE

CONCEPT AND USE CASE

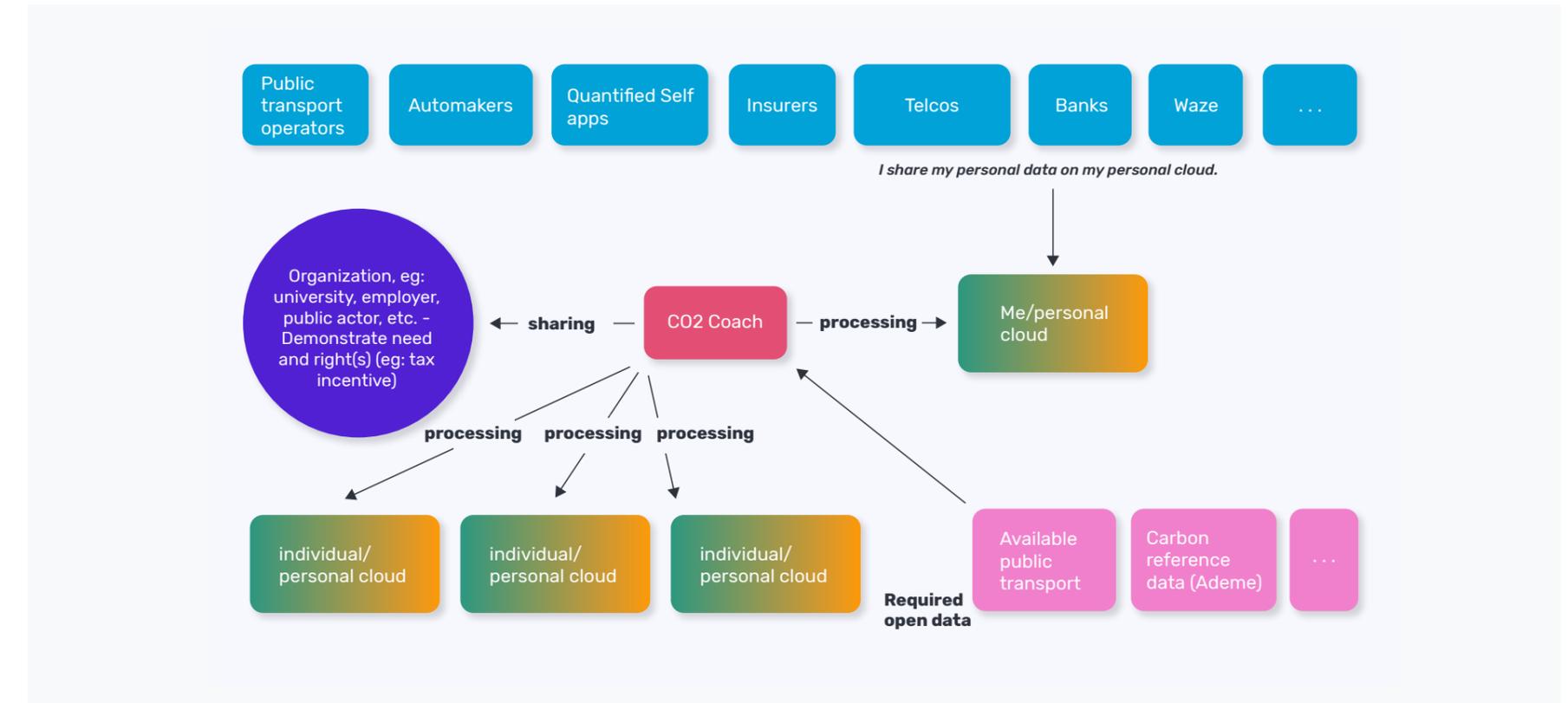


GOVERNANCE MODEL: PERSONAL CLOUD

"CO2 Coach" might be the first in a wide range of services. A lot of the personal data used to generate the service on people's personal clouds can also be reused by other types of mobility-related services (insurance comparison tools, public transport surveys, etc.). Users would not have to perform the task of importing their data multiple times to participate in several projects simultaneously.

To fully articulate the collective nature of the service – designed to enable colleagues, students, roommates and families to coordinate collaborative responses to common challenges and make collective requests, such as to modify public schedules in line with the majority, organize rideshares, and so on – will require 1) data sharing among CO2 Coach users' personal clouds,

and 2) data sharing from personal cloud to local public organizations, for example.

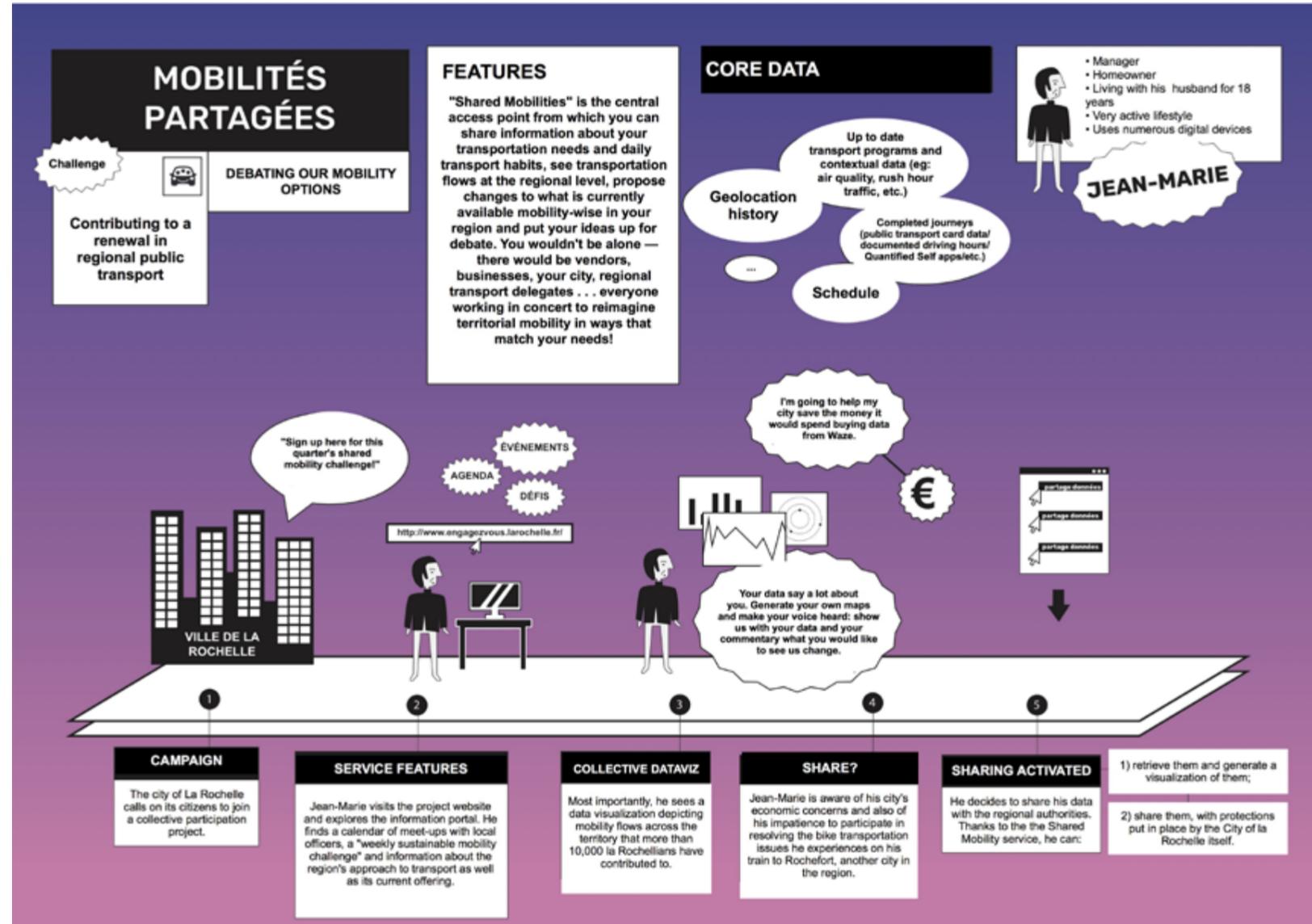


3

MOBILITÉS PARTAGÉES - SHARED MOBILITY

SELF DATA AND SUSTAINABLE MOBILITY - LA ROCHELLE

CONCEPT AND USE CASE

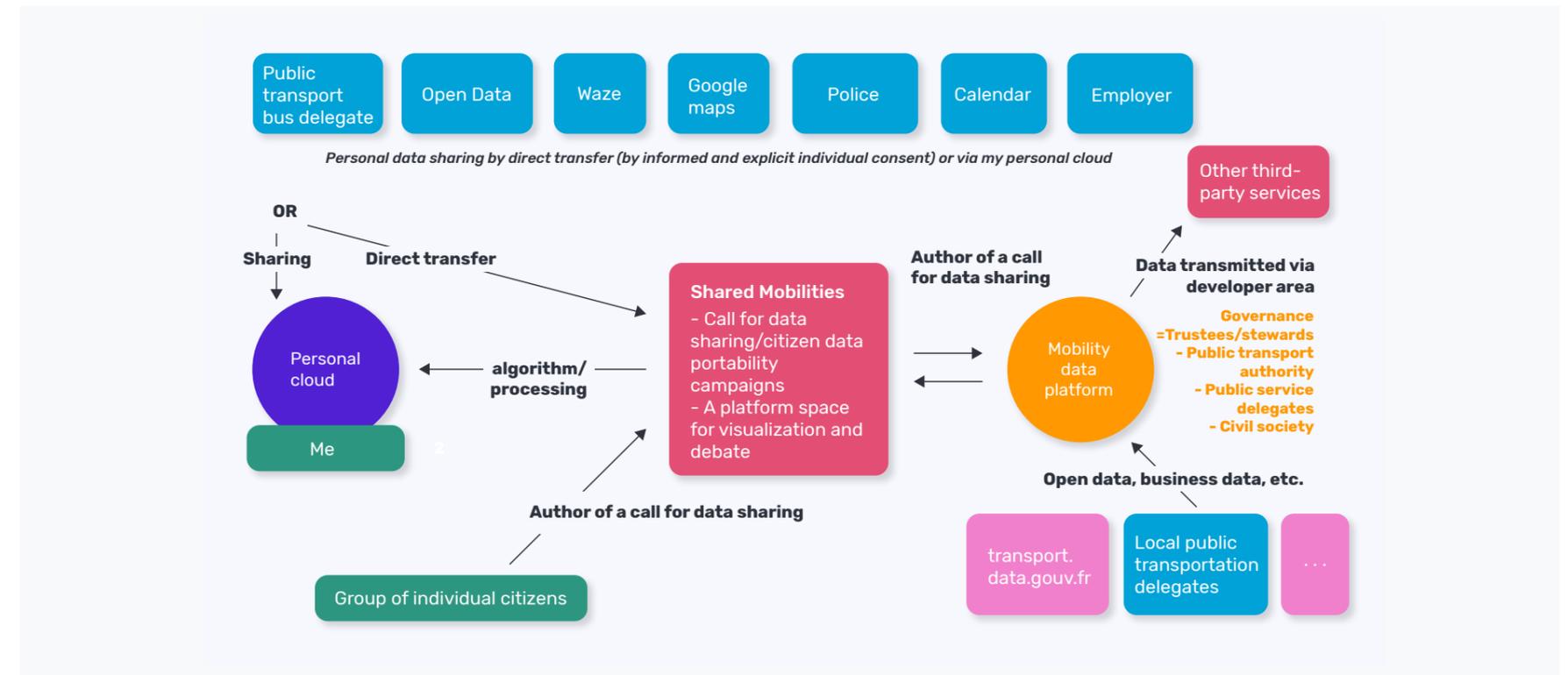


GOVERNANCE MODEL: CIVIC DATA TRUST

"Shared Mobilities" is not really a service, it's more of a communal space dedicated to calls for data sharing (and sharing facilitation) and citizen portability campaigns, plus a space for debate. It is backed by a public management model: the public actor is the guarantor that efforts are conducted in the general interest and has the mobility competence to ensure this (Public transit Authority). The Trust must nevertheless establish shared data

governance rules, and a shared architecture with mobility actors: those with whom the community has contractual links (Public Transport delegates) but also any territorial actors who are impacted by mobility changes (universities, hospitals, local shopkeepers, businesses) plus civil society at large (via social enterprise or citizens' jury). We hypothesize that direct data transfer and personal cloud sharing options would be the means

offered to individuals to access the services developed by reusers (via the developer area) and contribute to the Trust's data collection campaigns (eg: transport survey campaign). For developers, the interest in this model is that there is a relatively large group of individuals who would be using their services and the provision of rich public (and eventually private) datasets, albeit fictitious personal datasets.

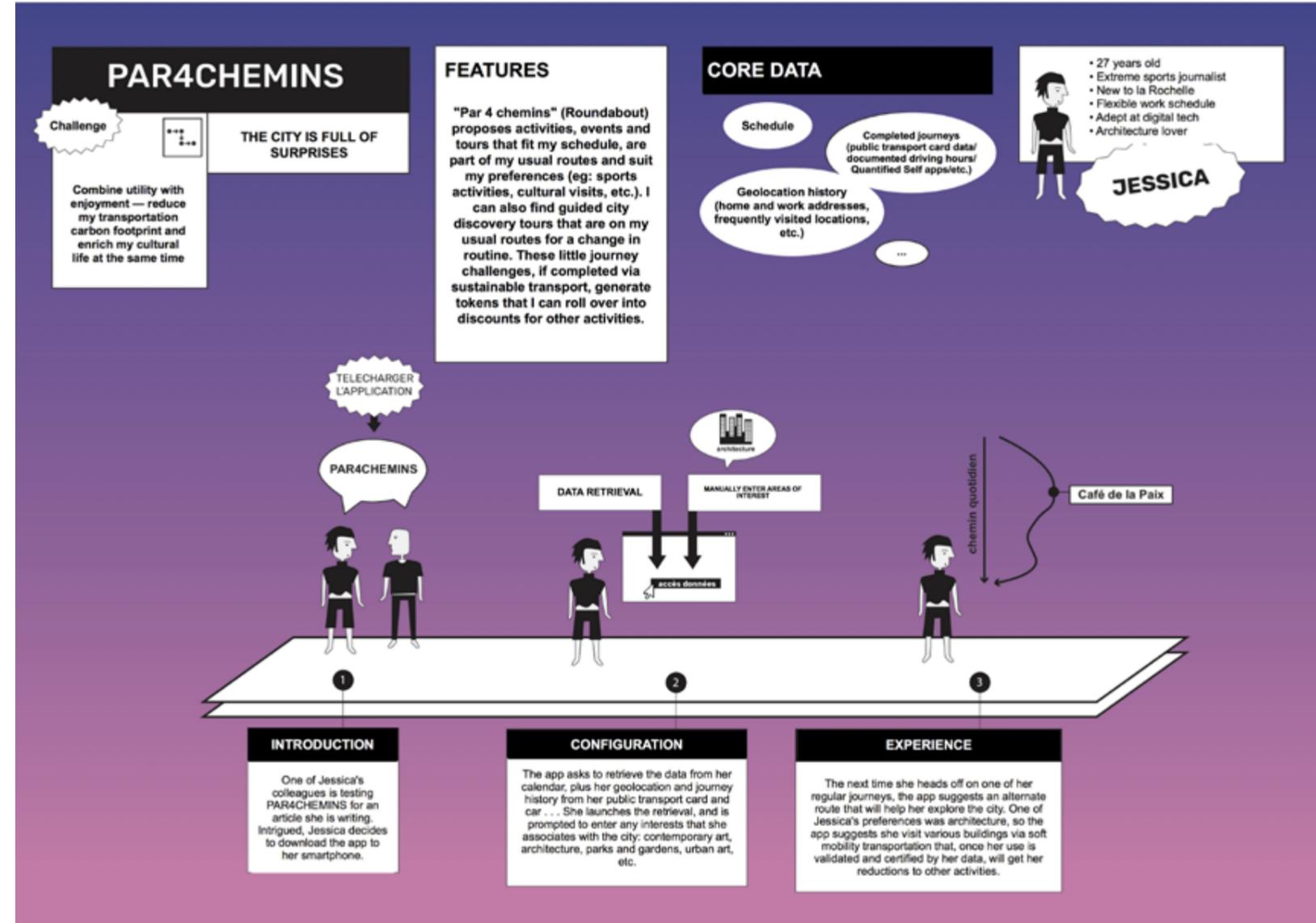


4

PAR 4 CHEMINS - ROUNDABOUT

SELF DATA AND SUSTAINABLE MOBILITY - LA ROCHELLE

CONCEPT AND USE CASE



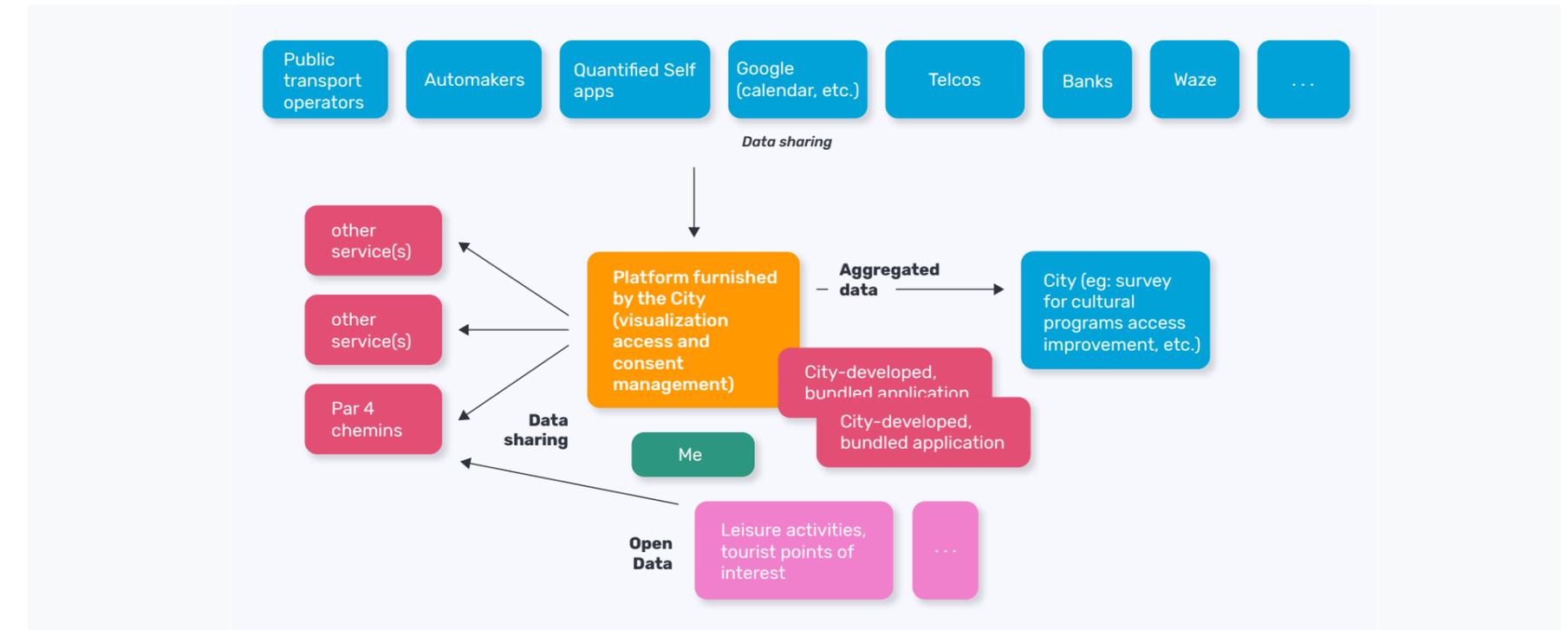
"Par 4 Chemins" (Roundabout) is a third party service built on the platform provided by the city. The city platform enables citizens to retrieve their data and store them securely.

It also has :

- 1) a set of city developed, pre-bundled services;
- 2) a list of third-party services certified by the city (APIs, developers area);
- 3) a management chart of consents granted.

When an individual decides to install "Par 4 chemins" – which is one of several third-party services that uses the personal data stored on the city platform – the service generates a list of the data the individual must share with service providers. Consent to personal data use is revocable at any time, and is easy to execute, and an effort has been made to allow individuals to carefully establish the data they wish to share. The city certifies trusted third parties after they have validated the

service's specifications and either established a common charter or signed a contract with them. A framework outlining which uses third parties can make of citizens' data has also been created: the city can not use the personal data stored on the platform, only aggregated data. In the long term, as far as collective uses go, short "campaigns" conducted via the platform could be effective (eg: "Contribute to a public transit survey with your data").

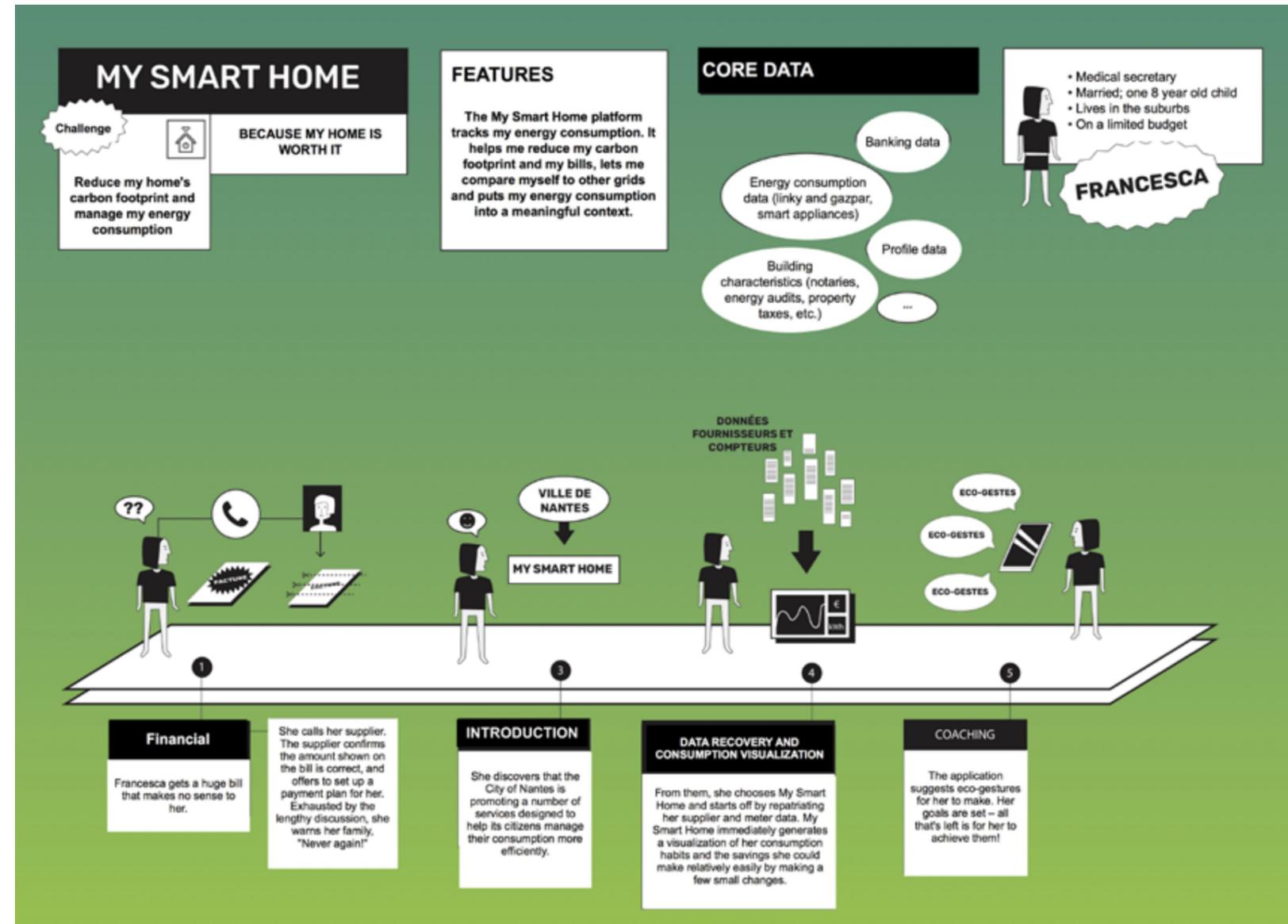


1

MY SMART HOME

SELF DATA AND THE ENERGY TRANSITION - NANTES MÉTROPOLE

CONCEPT AND USE CASE

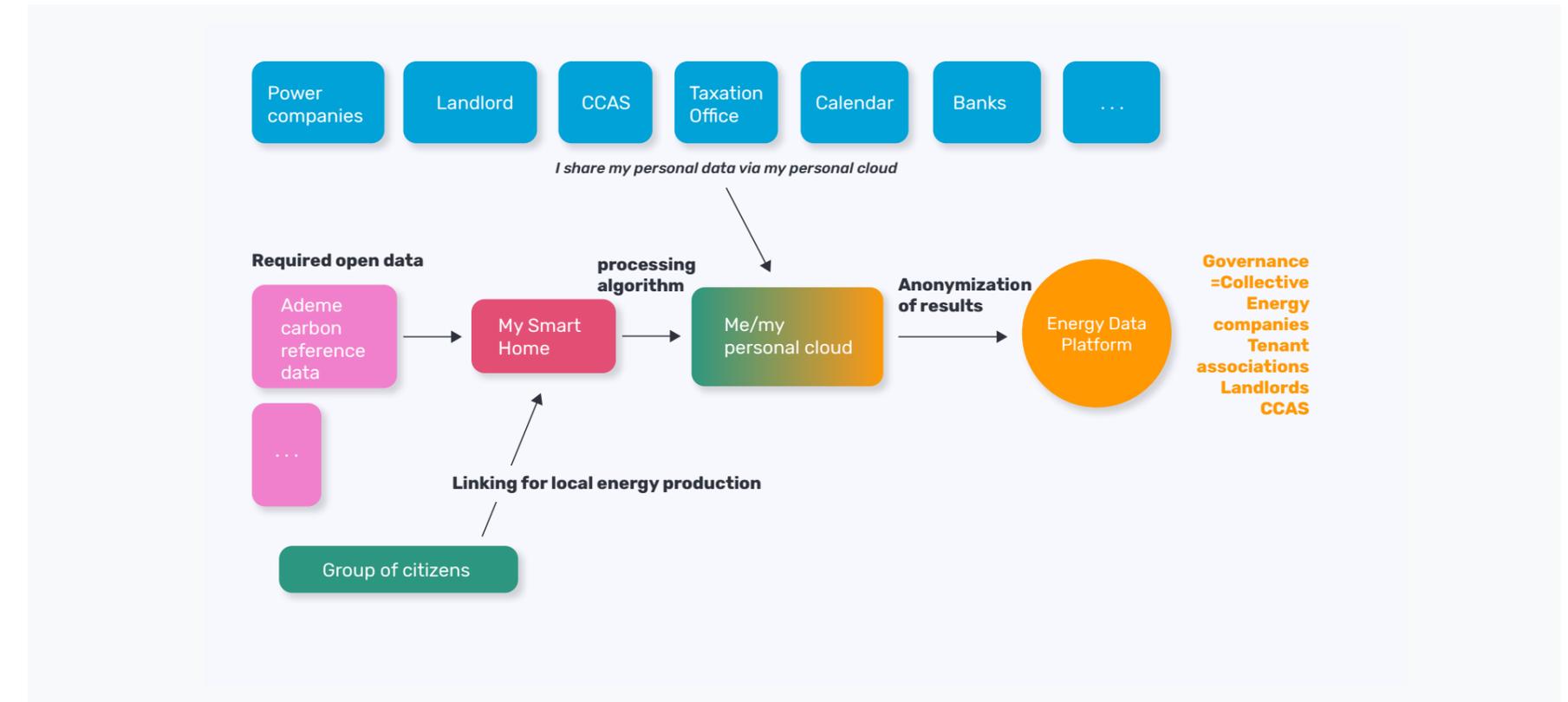


GOVERNANCE MODEL: PERSONAL CLOUD

My Smart Home is a personal cloud third party service. Despite their individualized nature, personal clouds also show great potential for collective use. To foster greater openness without distorting individuals' control over their data, establishing a Civic Data Trust is an option. It would be steered by a collective community using a mixed governance structure (eg: public interest group, local public

enterprise, joint syndicate) that includes the community and relevant stakeholders (tenants associations, CCAS, etc.). Its aim would be to ensure an open and fair use of data and build a framework of trust without blurring the lines between roles. The collective potential of the service should also not be overlooked: it offers a platform for users to obtain advice and make connections. It would,

for example, strengthen the power of tenants' associations by giving them a better grasp of their energy consumption. Finally, an appropriate model should also promote public social policies, such as those that seek to combat energy poverty.

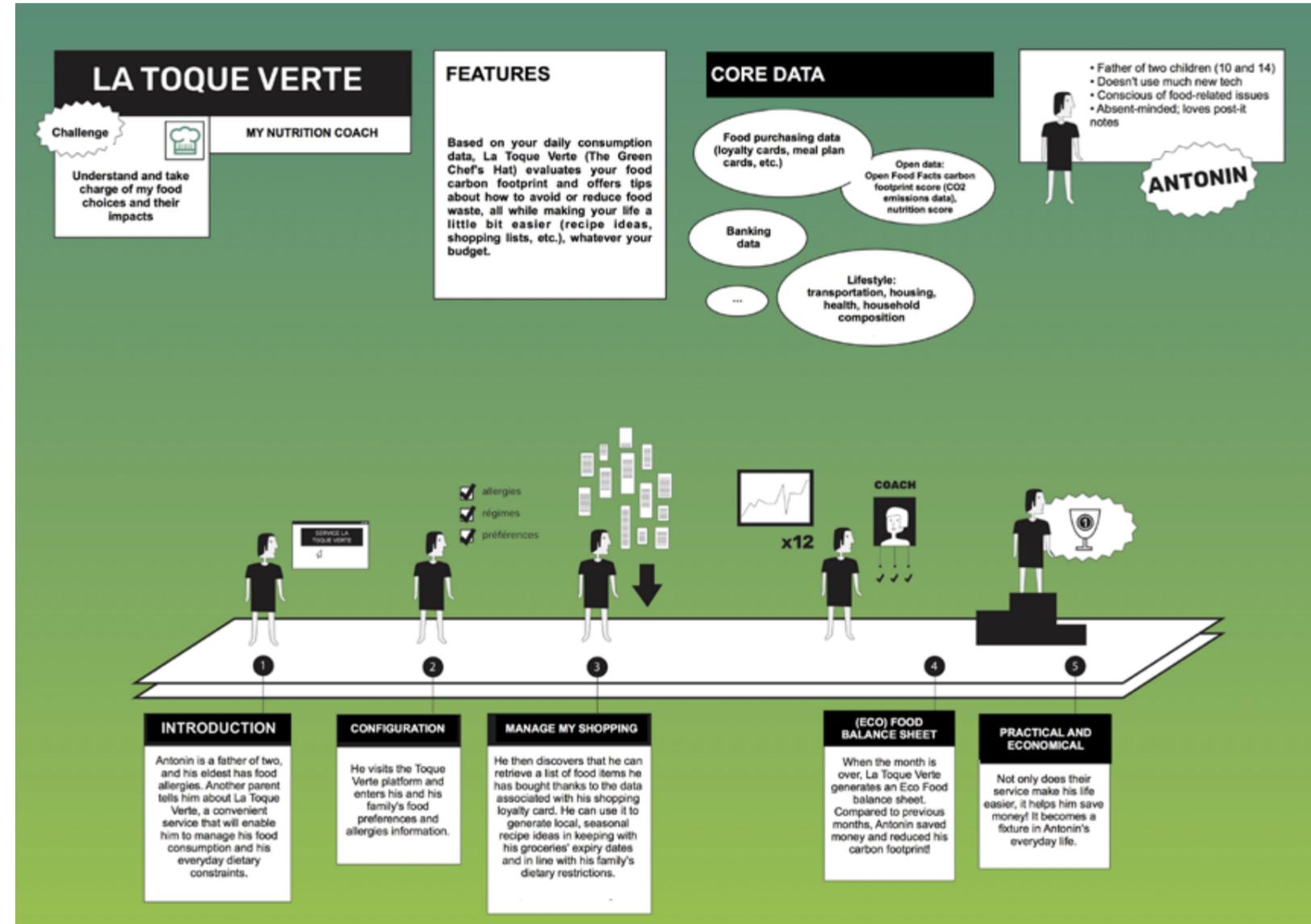


2

LA TOQUE VERTE - THE GREEN CHEF'S HAT

SELF DATA AND THE ENERGY TRANSITION - NANTES MÉTROPOLE

CONCEPT AND USE CASE

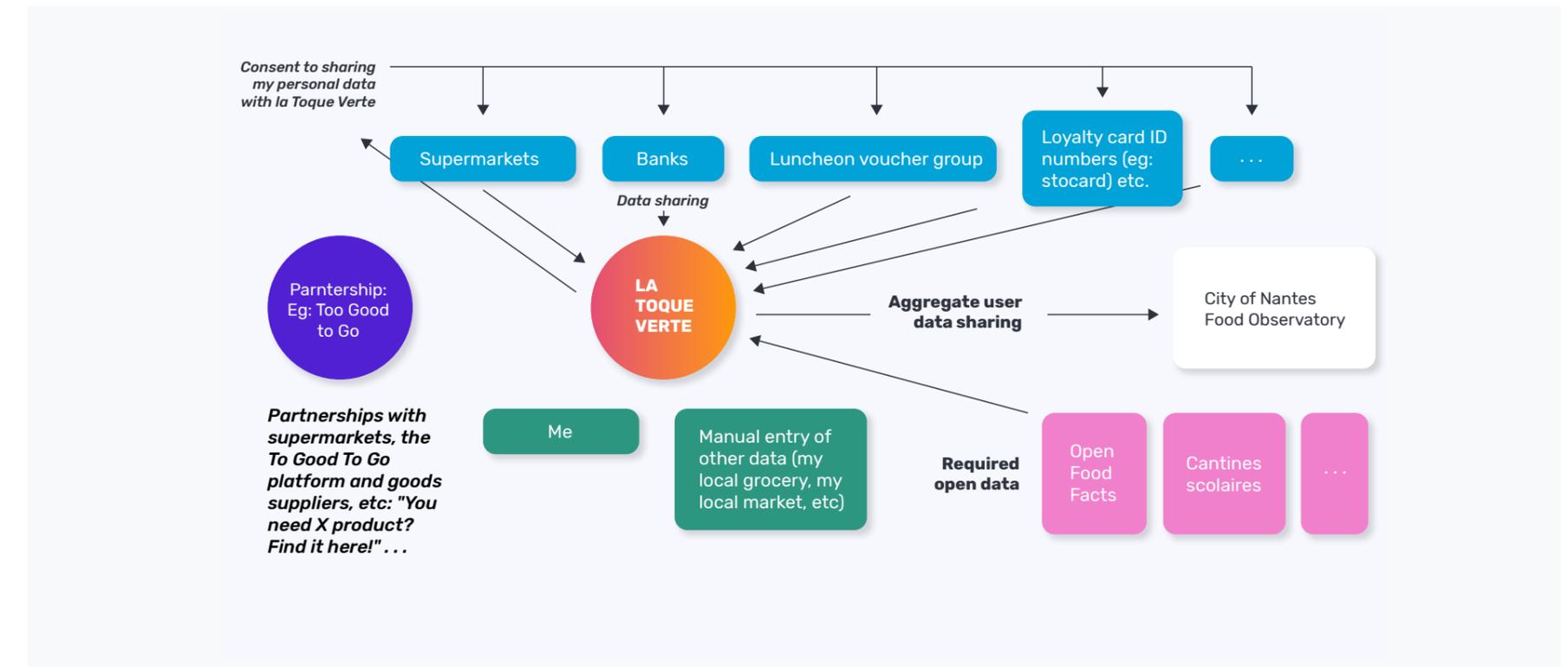


GOVERNANCE MODEL: DIRECT DATA TRANSFER

The Toque Verte is a very good example of the "direct transfer" model. The City of Nantes can launch the service even if only a small number of data holders – such as supermarkets and luncheon voucher groups – agree to share the data they hold with their clients. They do not need to provide a 360° overview of the data for the service to work, so the simplicity of a model based on explicit individual would likely suffice. Should

the number of data holders required to provide the service increase to include banks, or if the service decides to incorporate individual health data (food allergies, etc.), a transition to the "Personal Cloud" model would probably be recommended. This type of use case could be integrated with the city's Projet d'Alimentation Territoriale (PAT, Territorial Food Project) and the service could be constructed together with its citizenry. Working

with the future Sustainable Food Observatory would enable the city to use Self Data to create knowledge at the territorial level, thereby deriving the maximum benefit from citizens' personal data. Next step(s): encouraging local vendors to form groups and sign a common charter.

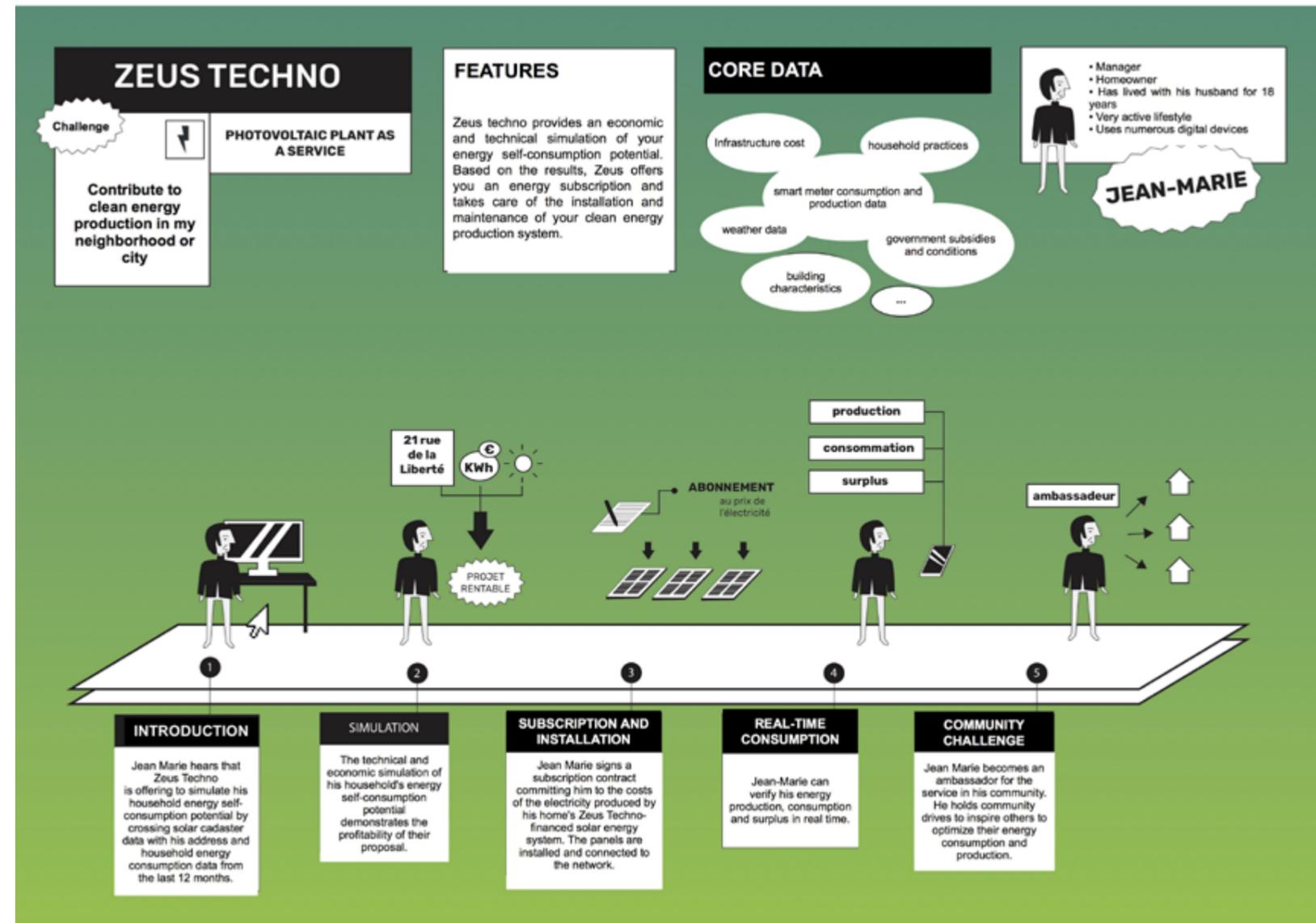


3

ZEUS TECHNO

SELF DATA AND THE ENERGY TRANSITION - NANTES MÉTROPOLE

CONCEPT AND USE CASE



GOVERNANCE MODEL: TRUSTED THIRD PARTY PLATFORM

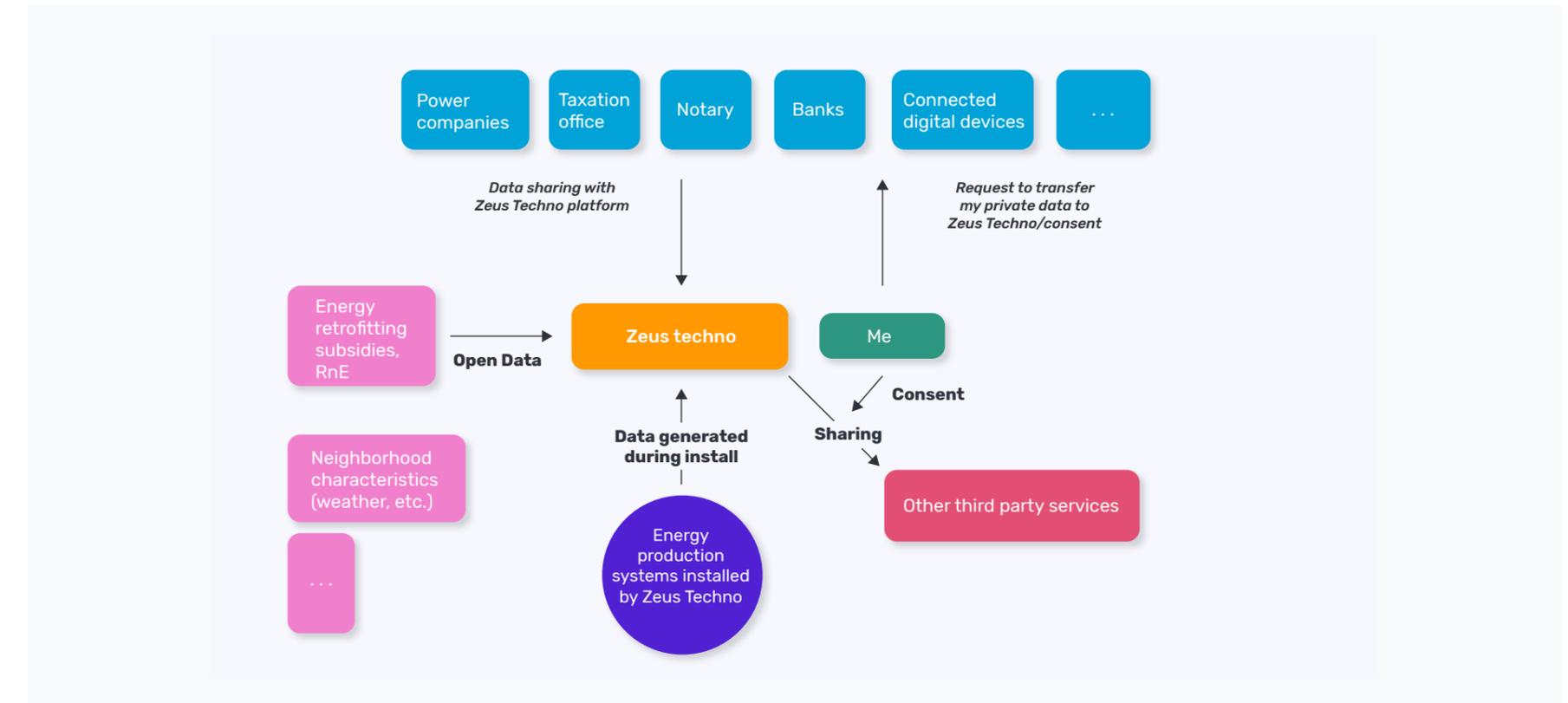
“Zeus Techno” is

» 1) An energy provider: they install solar panels, connect users to the network, and offer a subscription plan;

» 2) a service that allows individuals to

- a) simulate their self-consumption potential on the basis of their data and make “Zeus Techno” an offer; and
- b) manage their energy consumption, self-production and resale data (generated by Zeus Techno equipment); and

» 3) a platform: Zeus Techno has an API, which lets individuals share their data with other third-party services based on Zeus Techno’s offering if they wish, such as a recharge station for an electric vehicle, vehicle rental, etc.

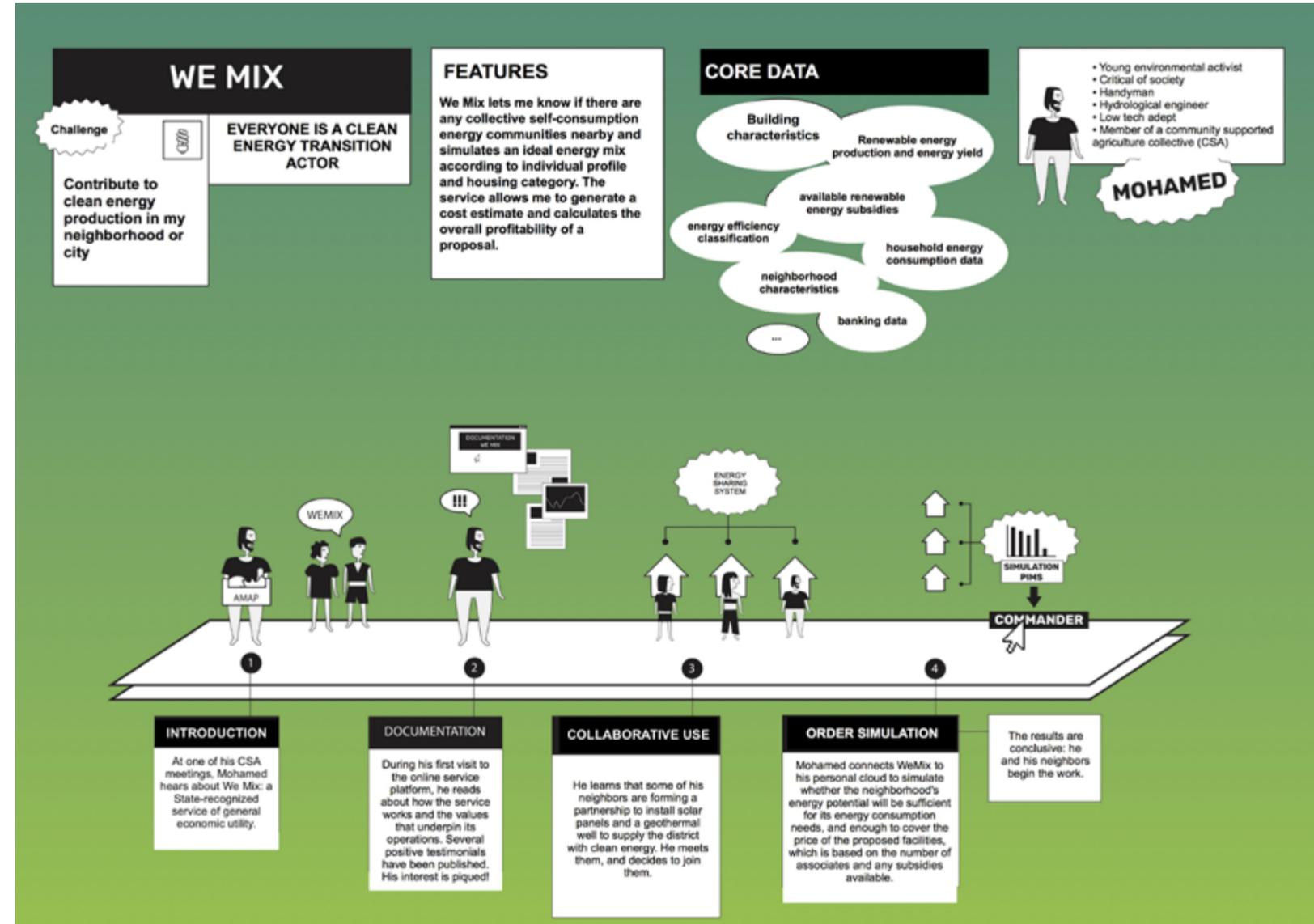




WEMIX

SELF DATA AND THE ENERGY TRANSITION - NANTES MÉTROPOLE

CONCEPT AND USE CASE

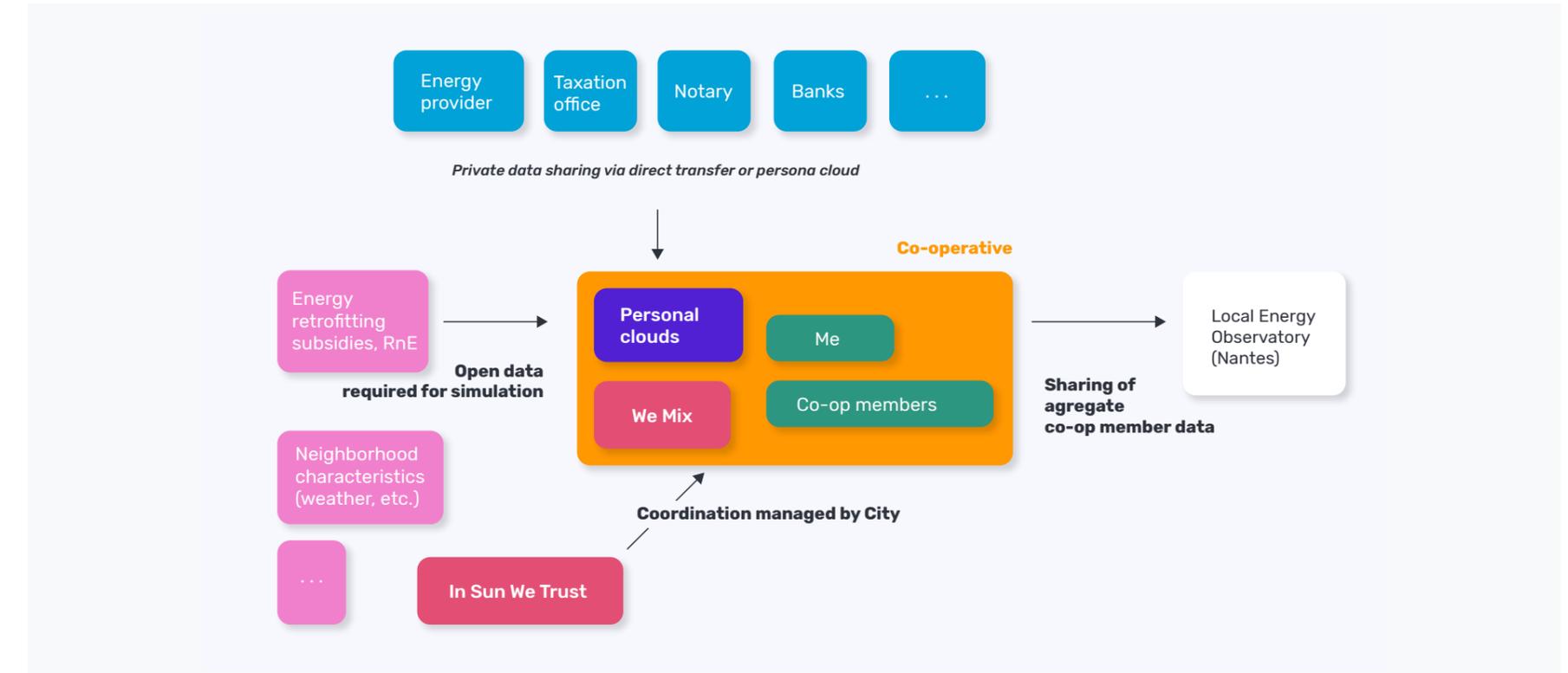


GOVERNANCE MODEL: DATA CO-OPERATIVE

We Mix is a data cooperative that manages an energy sharing cooperative. When combined, co-op member data establishes a critical mass of local citizens, enabling them to collectively build clean energy production facilities. It shares commonalities with a production co-op and an energy co-op. The members collectively decide upon (and share) the costs of infrastructure construction, production and maintenance. Data processing is central

to a proper needs and cost analysis: understanding which energy mix is appropriate according to members' profile data plus gaining a fuller picture of household priorities and their neighborhoods' environmental characteristics, such as the weather. Processing is also needed to manage the energy network (Smart Grid). WeMix ought to rely on free and open software. Achieving critical mass will make it possible to pay a salary to those ta-

king care of the electricity network. Calling on other regional actors such as FabLabs would make it possible to construct hardware such as mini-wind turbines. The City would be expected to integrate the initiative with other projects (ex: solar cadaster In Sun We Trust).



IN PRACTICE: OUTLINE AN EXPERIMENTAL SCENARIO FOR YOUR REGION

You now have a clearer idea of the kinds of data you can share with people and some concepts for uses that might emerge from sharing them. You have also started to think about governance models you might implement, and after the iterative workshops with your local collaborators, you know which ones you think you should direct your efforts towards. Ideally, you would have used the workshops to begin discussing the means you will use to put your experiment in motion.

The time has come to ask yourself a few questions before the next workshop, when everyone will come together to consider one or more experimental scenarios. In your view, what are the main objectives you wish to achieve with your Self Data Cities experiment? Who do you want to test your concepts, and how many testers would you need? Which use cases do you really want to try out?

By establishing these three starting points (objectives, testers, use cases) in advance, you won't start the workshop empty handed. However, you will want to further validate your objectives with the other workshop participants so that the interests

of your (potential) experiment partners converge with the Self Data Cities objectives and with the interests of the communities you will be serving. For example, you could ask them: "What would you like to learn/achieve/prove via a [theme] Self Data Cities experiment?" "What do you need for this experiment to be approved/your participation to be approved?" (eg: a clear legal framework that spells out exactly where the data will go, and when, and how); or "What do we want to avoid at all costs?" eg: an experiment launch date in three years, disinterested testers, etc.

Then it will be time to get together and define the experiment as it might exist, including its projected timeline, personal data sources and types, testers, actors, use cases/their development

In La Rochelle for example, we worked on two scenarios: "Shared Mobility" and "My Mobility Coach." Nantes Métropole went with "La Toque Verte," and envisioned a scenario grouping two use cases (My Smart Home and Zeus Techno) together, which they dubbed "Hestia & Zeus." In Greater Lyon, the chosen scenario was "My Social Journey," uniting two use cases: "My Direct Benefits" and the "Sortia Lyon."

In the following pages you will find a summary of each of these scenarios. The methodology that enabled them to emerge can be found in Appendix 5.

What follows are not the final roadmaps the cities adopted – after the last workshop, a succession of iterations both internally and with potential project partners allowed each of them to recalibrate their objectives, deadlines, timelines, and budgets. Preparing to launch an experiment before its official announcement takes time, especially for public actors who are constrained by electoral schedules, so be patient – and keep your eyes peeled, because these three cities are nowhere near finished with Self Data!

SELF DATA AND SOCIAL WELFARE – GREATER LYON

1. The "Mon Parcours Social - My Social Benefits Dashboard" Experiment

Process summary:

» 50 volunteers retrieve their social welfare/benefits status data from a personal cloud provided by Greater Lyon. They are all seeking to rejoin the workforce, and each is receiving the RSA (Revenu de Solidarité Active, a return to work welfare benefit) and so would greatly benefit from an overall picture of what is a rapidly and continually evolving situation -> Develop the necessary connectors between the testers' personal clouds and any data holders – Social Services, the city, perhaps the Tax Office, etc. – willing to share a copy of testers' data with them.

» Each personal cloud has an app "store" containing a city-developed app called "Ma Situation Sociale (My Social Benefits Dashboard)," from where they can get an overall picture of their benefits status, simulate benefits applications and pre-fill application forms with the "Mes Aides (My Benefits)" service (via questionnaire), and prepare a completed application file for submission to the appropriate department (by generating a "batch file" PDF combined with the data stored in their personal cloud). They share the file with their social worker, who has a restricted view of the dashboard ("TS interface"). This enables the social worker to get a better grasp of the tester's status, and provide any extra information needed for

the dashboard to generate an up to the minute case status summary sheet for the tester.

» Strong emphasis is placed on prompting testers and social workers to remember the app, visit the personal cloud, and utilize the tool #mediators #MDM (Maisons de la Métropole).

2. The "Sortia Lyon" Experiment

Aims:

» Simplify testers' access to discounted rates for cultural outings offered by the city.

» Better grasp the stigmatizing effect of presenting reduced rate vouchers and invent ways to avoid it.

Process summary:

The "Sortia Lyon" service is derived from the "My Social Benefits Dashboard," and it uses the same data. It can operate as a standalone service based on testers' calendar/budget preferences, or it can work in tandem with the Dashboard. When the two are deployed together, three additional concerns/recommendations arise:

» During the experiment prep phase: an exploratory sociological study will enable you to clarify users' needs in greater detail. What are the stakes surrounding the social stigma attached to rate reductions? What are the barriers to asking for and obtaining them? Who are the target audiences?

» Potential testers may be different from those using the "My Social Benefits Dashboard." You may wish to conduct a separate recruitment process based on the findings from the research mentioned above.

» "Sortia Lyon" could also be developed as a dashboard that gives an overview of the user's social welfare profile and a list of cultural activities available in Lyon according to their preferences, including the preferential rates they are entitled to. They could click on an event they like, visit the relevant page, and either buy the ticket by furnishing the necessary proof of their reduced rate entitlement directly from their personal cloud, or buy a ticket at the reduced rate and generate a PDF/data printout that they must present to gain admission to the event.

SELF DATA AND SUSTAINABLE MOBILITY - LA ROCHELLE

1 - The “Mobilités Partagées - Shared Mobilities” Experiment

Aims:

- » Validate the intuition that shared personal data can practically contribute to city management, common knowledge gathering, etc.
- » Experiment with the factors affecting the individual choice to share personal data to serve the common good.
- » Explore the potential of data co-operatives.

Process summary:

- » 50 testers (Greater Lyon (CDA) employees or postal workers; both are categories of civil servants) are given access to “Traces,” an open source app (built by the Fabrique des Mobilités de l’Ademe), which enables them to retrieve their personal geolocation data and view a dashboard summarizing their journey data (statistics visualization, CO2 footprint calculation, etc.) from which they can share these data.
- » At least two other types of data are shared on Traces: CDA Yélo app data (public transport, public bicycles, and self-service cars, etc.) and, if possible, Thalès (parking space data). -> Connectors are developed to allow testers’ CDA

and Thalès data to be transferred to Traces.

- » Testers collectively agree to share their data for research etc. during meetings, debates (using a co-op structure: one person = one vote) -> There is no way to obtain significant findings with only 50 testers, so the goal is to study individuals’ motivation for sharing their data.

2 - The “Mon Coach Mobilité - My Mobility Coach” Experiment

Aims:

- » Test the Self Data hypothesis: “sharing my data gives me more control over them.”
- » Determine whether it is possible for testers to obtain a useful/practical measurement of their carbon footprint/mobility budget based on the data shared with them.
- » Observe behavioral changes (and in parallel, ascertain whether the use case can be developed further).

Process summary:

- » 50 CDA employee or postal worker testers are each allocated a personal cloud enabling them to manage their data.
- » Their Yelo data (public transport, public bicycles, and self-service cars, etc.) and potentially their Thalès (parking use) and SNCF (train tickets) data are shared with them.

-> Development of data holder connectors to the personal clouds.

- » An application developed by a CDA provider is available on the cloud’s app “store” that generates a visualization of their carbon footprint and mobility cost/consumption data, offers them personalized advice about how to reduce both (it remains to be seen whether it is financially feasible to develop this feature), allows them to share their CO2 results score with the CDA’s carbon aggregator. -> A specific app to be developed. Testers to be overseen and supported (by a dedicated researcher) for the duration of the experiment.

SELF DATA AND THE ENERGY TRANSITION - NANTES MÉTROPOLE

1. The “Hestia & Zeus” Experiment

Aims :

- » Estimate if access to and reuse of their regular consumption data and cross-referencing it with others allows individuals to reduce their consumption or encourages them to commence energy retrofitting of their homes.
- » Understand how digital support provisions like Self Data can be developed to effectively deploy targeted, tailor-made citizen challenges (eg: a “positive energy family” challenge), and by association, make it possible to reach a wider audience.

Process summary:

- » 100 testers form the 2020/2021 cohort of Positive Energy Families: half of the panel has already had their home retrofitted, the other half has not and are still living in non-energy efficient homes. A second panel of 1000 testers, who are not necessarily homeowners, is recruited, but will receive less face-to-face support.
- » Members of both panels are allocated a secure personal space allowing them to retrieve their data from partner holders (Enedis, GRDF, EDF, Engie, etc.) and to add data concerning the characteristics of their housing and the dates for/invoicing details from their renovation work.

-> Public empowerment – the hope that offering dashboard type services, renovation assistance, etc. will sufficiently enrich existing services rather than needing to develop a service from scratch.

- » The testers will be able to reuse their personal data in different ways. Step 1 “HESTIA:” generates a data visualization and a series of targeted eco-challenges that result in more personalized advice about energy efficiency measures once achieved. Step 2 “ZEUS:” predicts the ROI from energy retrofitting work by owners who have already completed the work and visualizes the savings to be made by owners with similar profiles to raise awareness about the advantages of retrofitting over time.

2. The “Toque Verte - Green Chef’s Hat” Experiment

Aims:

- » Help people better understand the carbon footprint they create from their food consumption, and give them a simple and concrete overview of the impact that their purchases make.
- » Offer individuals the opportunity to choose alternatives that are more respectful of the environment.
- » Contribute to the creation of common knowledge – allow people to share some of their (anonymized, pseudonymized) data with France’s Sustainable Food Observatory.

Process summary:

- » 100 testers form the 2020/2021 cohort for the Family Zero Gaspi (Zero Waste) Challenge (face-to-face coaching, etc.). They are equipped with an application that retrieves their personal food data from at least three data-holding partners (supermarkets) so they can monitor and reduce the carbon footprint created by their food consumption and increase their awareness about food waste and the impact of their purchases.
- » A recruitment campaign is underway, conducted by Nantes Métropole to recruit experiment volunteers. They must be customers of at least two of the three data holders who have agreed to share data during the experiment (ie: they must have a loyalty card). -> Development of data sharing connectors.
- » The testers will then be able to use their personal data to visualize their food’s carbon footprint; find alternatives for certain purchases; contribute to the creation of common knowledge; compare themselves with others; and share their experiences. -> App development – calculation of the carbon footprint crossed with Open Food Facts data + Ade-me Base Carbone®.



04

**“SHALL WE BEGIN?” –
PROTIPS FROM EXPERT
SELF DATA ACTORS**

04

“SHALL WE BEGIN?” — PROTIPS FROM EXPERT SELF DATA ACTORS

Nantes Métropole, Greater Lyon, and La Rochelle are Self Data pioneers. Each has managed to foster the emergence of a dynamic in their metropolitan areas by putting together a group of motivated actors and ambassadors, having a precise idea of the data that are available and several use cases that put these data to concrete use, moving towards appropriate governance models, and creating solid experimentation scenarios that pave the way forward. The meetings attended by all three cities were crucial to their ability to structure the projects and iterate potential opportunities and solutions that adopting and implementing such policies might entail. Data sharing is one of many areas where Self Data Cities may need to cooperate, as certain data holders are common to different cities and also different themes. Coordinating would facilitate experiment implementation across the board. If you are embarking on a Self Data Cities experiment, we hope that you will share your strategies and intentions with collectives and coalitions of like-minded others!

A Self Data experiment has three essential ingredients:

- » **testers: individuals who will control their data during the experiment, learn from the practice, and possibly be the subjects of observation by researchers;**
- » **personal data “returned” by data holders: per-**

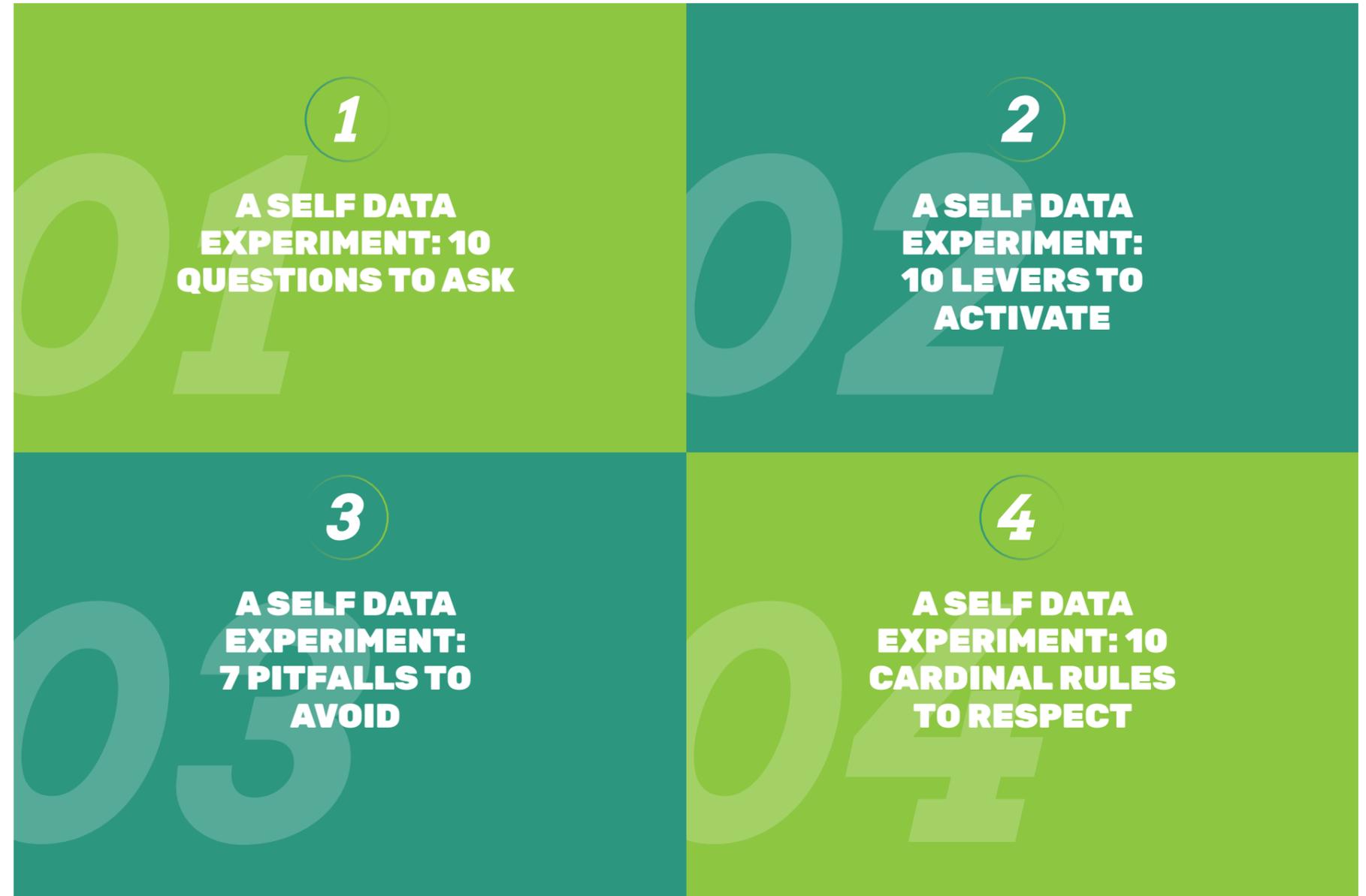
sonal and user data-holding organizations set up transmission channels to share these data only with the individuals concerned;

- » **uses for these data: tools developed to allow testers to derive value from their data (store them securely, manage them, share them, cross them with other datasets, and reuse them).**

Bear in mind that, for the purposes of this kit, the following tips refer to fully fledged Self Data experiments and not to pilots. The objective of an experiment is to learn, even if it means making mistakes, while the goal of a pilot is to demonstrate the viability of a concept. A Self Data experiment is thus a conjoining of three types of experiment:

- » **a sociological study;**
- » **an experiment in driving change in large organizations;**
- » **an experiment in innovation policy implementation (launching new tools).**

To guide your efforts, we have put together a list of pointers to help you conduct what may be construed as a slightly crazy experiment with Self Data.



1

A SELF DATA EXPERIMENT: 10 QUESTIONS TO ASK

01 WHAT ARE THE OBJECTIVES OF THE EXPERIMENT? WHAT DO WE WANT TO TEST?

It is important to identify your main objective: having your fingers in too many pies will inevitably become problematic. Formulating and not straying from the experiment's primary objective is essential to achieving its goal (to demonstrate the value inherent in Self Data). However, despite the backbone it provides during the experiment's preliminary phases, you must reconsider it once you have compiled your final roster of contributing actors. The variety of domains you will traverse during the experimental process will add many secondary objectives to the pot (for example, "scaling Self Data"), which must remain subordinate to your primary aim.

02 WHICH DATA WILL BE SHARED WITH THE TESTERS? WHO CURRENTLY HOLDS THESE DATA?

A Self Data experiment without data holders is like a fondue without cheese — it makes no sense. You will have to surround yourself with holders from the start and work alongside them to set up the sharing channels. You will need to find internal ambassadors at each who will champion the drive internally to validate the strategy and bring about its implementation.

03 WHO ARE THE TESTERS? HOW WILL WE RECRUIT THEM?

This question is bound to create tension. Who should you recruit as testers: should they be early adopters, or at least people who are already interested in learning to manage their data? Or maybe people who aren't comfortable with digital technology and could benefit from being introduced to it? Choices will need to be made.

04 WHICH LOCAL ACTORS DO YOU WISH TO ENLIST? HOW DO YOU PLAN TO WORK TOGETHER WITH THEM?

Between data holders, social enterprises, SMEs, competitive clusters, and researchers, you will bring together a wide range of stakeholders whose needs and strategies might actually conflict. Remember to define how this consortium will work, perhaps by drawing up an official agreement (rules of communication, entry into the consortium, decision making, etc.) that acts as a charter that these actors commit to following — and also a charter that explicitly defines the research scope/user feedback protocol.

05 WHAT SHOULD TESTERS BE ABLE TO DO WITH THEIR DATA?

You will not immediately have the answer to this question, which is normal. It brings back the old Self Data dilemma between which comes first, data or services? It's like choosing between the chicken and the egg. Should you share the data first, and then see which uses emerge before you decide which data are the most appropriate? As part of the MesInfos pilot project, that is what we did — the goal was not to operate under any preconceived notions about the value of the data or its potential uses. But the risk of focusing our efforts on datasets that had little to do with each other was great. In the case of a Self Data Cities project, you should do things the other way around: you should know, before starting your experiment, which use cases are going to emerge — while, at the same time, maintaining a certain degree of flexibility — and therefore which data should be shared.

06 WHO DEVELOPS WHAT? HOW?

Any definition of a "Self Data Experiment" would be incomplete without mentioning third-party, data reuse services, which enable users to derive value from their data. "Third-party service" can also mean "data aggregation platform" (eg: personal cloud, third-party platform, data trust, etc.). The local public entities who pioneer Self Data initiatives do not need to personally develop their own platforms and services. They can rely on the local ecosystem, existing services, or even conduct open public tenders, etc.

07 WHO PAYS FOR WHAT?

Experimenting with Self Data is expensive: you will be paying project managers, technical service providers, designers, tester panel supervisors, etc. Once you have drawn up your budget, you will need to create a financial partnership orientation file that details the cost of participating in the experiment . . . and then find these partners. For them, you may wish to create a tiered partnership structure. You will also need associate project partners (eg: social enterprises, competitiveness clusters, etc.), because their expertise is essential. Financial partners are often data holding organizations. They may become partners for two reasons: either they intend to share their data during the experiment (and in this case, their financial commitment would have to be considerable, as they will require a considerable amount of support), or they wish to learn how to share personal data over the medium and long term.

08 WHO MANAGES THE PROJECT? HOW?

The project should be staffed by at least one (at least!) one full-time FTE, depending on the size of the experiment. The local public entity leading the experiment can delegate project management to another service provider, but that person cannot replace the city employees who act as Self Data ambassadors, as they will be the ones convincing data holders to join. You might also wish to break the project down into several moving parts — legal, technical, recruitment, tester management, and research, for example — that run simultaneously.

09 WHAT ARE THE FIELD TESTING INDICATORS? WHAT DO WE WANT TO LEARN?

This is the question that every large organization will ask you, especially if you have applied for national or European funding. Although you are here to learn, you can establish some indicators — not of success, but of change and evolution. For example, if your goal is to restore trust between individuals and organizations, measuring trust would become an explicit part of the research process.

10 WHICH TIMELINE TO FOLLOW? AND WHAT IS THE "REAL" TIMELINE?

The major phases of your experiment process, at each level, should be thoroughly mapped out: set the date when the entire cohort of testers will have the tools in hand, for example, and the date when data holders will be expected to have activated the final shared data file, etc. A Self Data experiment is complex — there are many actors, and they do not always make progress at the same pace. You will need several Plan Bs and an extremely robust decision making protocol, so that when a pillar of your project timeline map collapses, you can decide what to do as rapidly as possible.

2

A SELF DATA EXPERIMENT: 10 LEVERS TO ACTIVATE

If that list of questions worries you, never fear — Self Data experiments also come with one overarching benefit: everyone wins. Despite the profound changes that implementing Self Data entails, every actor involved — citizens, regions, data holders, digital service providers — is a player in the game, and each stands to gain something of value. We would even go so far as to say that Self Data is intrinsically positive and democratic. Stakeholders, regardless of their role, position, function or status, will (almost) always interpret “human-centered” as putting themselves at the center. This is an enticing prospect, and it should win you many allies.

01

GET CITIZEN/CIVIL SOCIETY ORGANIZATIONS AND CONSUMER GROUPS ON BOARD

The topic of personal data is now well known enough to have sparked the interest of civil society and consumer organizations to be interested in it, and even a desire to tackle its challenges. The participation of these groups is integral to the success of your project — how else do you expect to put individuals at the center of their data management without making them part of the process?

02

GET THE CNIL ON BOARD

Because Self Data embodies “the spirit” of the GDPR (and the 1978 Data Protection Act), the CNIL — mainly via its Innovation and Foresight Laboratory (LINC) — has been an integral part of Fing’s MesInfos pilot study and various Self Data experiments (including during the planning stages) from the start. Getting the CNIL’s stamp of approval (which means having the right to use its logo in your supporting documentation, if you are able to establish a partnership), even informally, will be necessary for you to inspire confidence in your stakeholders. It will be a decisive factor for future data holding partners. Fing’s main advantage is that we are independent from our partners, which means we occupy a space of “neutrality” during discussions. This greatly facilitates dialogue with representatives from our partners’ various departments as we form hypotheses and table risky solutions within the wider context of an experiment that is both audacious and also limited (in time, in participants).

03

DRAFT AND RATIFY A CHARTER FOR YOUR EXPERIMENT

A charter is a tool that establishes trust and creates transparency in your project communications. It is also a valuable internal tool that conveys both the project’s vision and its limitations over the coming months. It will enable you to state openly where the project’s priorities are as regards more contentious issues (for example the “data monetization” game) and being able to refer to it will free the experiment from any time-consuming controversies that might erupt. The charter also represents the very first collective output generated by the project actors, and the drafting and ratification process offers you the perfect opportunity to test the group’s dynamics and its decision making process. It will make your partner’s positions quite plain, and so enable you to identify solutions and compromises upstream of operational commitments.

04

FORM CLOSE TIES WITH YOUR PARTNER’S LEGAL SERVICES

Absolutely zero data will be shared without the sanction of your partners’ legal departments. It is essential that you gain their approval early on, as soon as you have settled on your experiment scenario.

05

PREPARE THE EXPERIMENT’S “ELEVATOR PITCH”

Having an easily understandable, basic description of the project will go a long way towards avoiding misunderstandings, and also might prevent misrepresentations of the project from appearing in the numerous press articles that will undoubtedly be written. It is a particularly indispensable tool given the number of actors involved and the innovative and transformative aspects of Self Data. You can bolster your pitch (or “the line of your argument”) with a few personal data reuse cases that clarify the subject matter and make whoever you are speaking to want to know more. In a similar vein, be sure to maintain a regularly updated FAQ so you have your most solid arguments and answers to hand — it will help you avoid being taken by surprise.

06

BEING “OPEN” OPENS DOORS

At the risk of sounding rhetorical, the philosophy behind open source software is perfectly in line with the values that underpin Self Data. Many even see open source as a prerequisite for establishing transparency in personal data processing. It can also lead to fruitful cooperation between projects, broadening the support and impact of your experiment. The open (and potentially free) aspect greatly simplifies the adoption and use of software solutions and tools for the project, especially compared to proprietary solutions that could require confidentiality agreements, contracts, etc.

07

IMAGINING DESIRABLE FUTURES IS FUN AND GRATIFYING — CAPITALIZE ON THAT

Imagine a “Siri” that knows your favorite pizza (but doesn’t sell the info to Pizza Hut); a GPS that could tell you not only how to reach a destination, but also which destination would be the most relevant; or a fridge that reprimands your teens when they drink too much soda . . . Any use is possible in the world of Self Data, and because no data is shared without being under the control of the individual, concerns about tracking or surveillance can be set aside when dreaming up new uses. This allows your collective to embark on a quest for more innovative and unpredictable approaches. Imagining new services — be they useful or “useless” — and projecting yourself into a future where they are possible and real can be very fun. Creating imaginaries creates energy, enthusiasm, and cohesion within the group.

08

BREAK DOWN SILOS WITHIN ORGANIZATIONS

Personal data within organizations is transversal in nature. So it will naturally transform your Self Data experiment into an inter-departmental training ground for each of your project partners. The value generated by this type of interchange is tangible, and has the merit of being more easily recognizable than the virtues afforded by Self Data. The practicality it requires is useful, and can serve as a lever for you to convince your partners to make themselves more available for the experiment.

09

USE PEOPLE’S PERSONAL COMMITMENT AS LEVERAGE

As mentioned above, Self Data values appeal to collaborators and project actors as individuals, as citizens, as family and community members, and so on. Propagating Self Data values is motivational, and fosters close ties among its proponents. This is an incredibly useful lever: it will help you not only overcome the many obstacles that the project is bound to encounter, but also spread the word and sustain the vision of Self Data over the longer term. This alchemy is something you must cultivate through general assemblies and events where different actors can meet and exchange. These may give way, later on, to more informal gatherings. The more they can get on with things without you, the better the project will be!

10

GET THE INTERNATIONAL MYDATA NETWORK BEHIND YOU

The network will be delighted to hear about your decision to launch a concrete initiative (remember: MyData = Self Data) and will promote your efforts among network members — these global experts and international organizations have much experience and insight to offer you.

A SELF DATA EXPERIMENT: 7 PITFALLS TO AVOID

Here are some of what we now think of as “bad good ideas” that can appear during your experiment. These are the things we wish we had known when we started the MesInfos pilot study and experiments.

01 HACKATHONS: DON'T DO THEM (?)

This rallying cry is a recipe for how to best serve your purposes in the most pragmatic way possible. Hackathons and app development competitions cannot generate proof of concept (reliable, easy to use, intuitive, etc.). You can use them to spread the word about new opportunities and to inform project development, and they can help you push ideas a step further, create buzz around a topic, recruit new team members. But the question remains: will any of that help you achieve your goals?

For example, if the primary goal of your Self Data experiment is to promote the concept and eventually embed Self Data values in the digital services of tomorrow, then a hackathon is an interesting tool. On the other hand, if the objective of your experiment is to observe how individuals use a Self Data service, then you will need to avail yourselves of a service that is already reliable enough — because it has been developed more fully — to be used autonomously by individual citizens.

02 ANTICIPATE SLOW PERIODS IN TESTER ACTIVITY

Mid-August (in Europe), the week between Christmas and New Year's Day, months when national holidays come one after the other, back to school time, school holidays... At certain times of the year, it is nearly impossible to anticipate how available your testers will be, especially since that availability also depends on the use case: on Christmas day, a smartphone game will have its highest audience, and a research questionnaire its lowest response rate. If the subject is complex, experienced tester community facilitators will likely be aware of these constraints.

03 EXPECT TESTERS TO LOSE INTEREST

When it comes to experiments that go on for months, the majority of the testers recruited at the beginning almost inevitably give up at some point during the experiment. In any case, that's what the statistics say. Retention and drop-out rates are well-documented features that must be taken into account when planning the technical side of sociological research and should be factored into any goals for numbers of recruits. You must also consider recurrence of use vs. experiment duration: if your use case is intended to allow individuals to move houses more easily, or to change a service provider, remember that very few of them will actually move house, and that switching providers might only happen one time during the entire life cycle of the experiment.

04 START RECRUITING ONCE THE TESTING ENVIRONMENT IS READY. .. AND NOT BEFORE

Avoid making your tester panel wait! If you do, your panel will certainly be let down in some way, and this will accentuate any desire they have to disengage from the proceedings — all of which means that you run the risk of conducting the core experiment with a drastically reduced number of testers.

This pitfall is, unfortunately, very tricky to avoid: recruitment is a lengthy process, and one that often involves partners (who also have their schedules and priorities); the data are always “nearly ready;” the solution is always touted as being ready “this time” . . . all the while, the experiment's hard deadline is inexorably approaching . . . and we have not even begun to list the external factors that will weigh in! The reasonable solution? See numbers 5 and 6 on this list.

05 JUGGLING DREAMS AND REALITY

A Self Data experiment is an opportunity to envisage what might be possible when conditions are favorable. Creating this imaginary is crucial to clearly identifying and consolidating your vision. But these concepts are impossible to achieve within the experiment's time frame, and with existing means. You must learn to juggle enthusiasm and expectations for particular outcomes with clarity regarding the experiment's real perimeter of possibilities. It's up to you to find the balance that will allow your experiment to succeed.

06 HAVE A PLAN B FOR EVERY MAJOR STAGE OF THE PROJECT

As noted previously, a Self Data experiment turns into a succession of experiments: in change management inside large organizations; in launching new software; in gathering results for a sociological study. Even though each produces new knowledge, none will achieve precisely the results you anticipate, which destabilizes proceedings downstream. To ensure that the last experiment in the chain takes place under the appropriate conditions, you will need to ground it on a solid foundation — some practical outcome that is perhaps not pioneering. Aiming for a less ambitious but proven solution will temporarily bridge gaps in your timeline and enable you begin the next phase of the experiment at the scheduled time.

07 GIVE YOURSELF LEEWAY IN YOUR SCHEDULE

To be on time, you must plan for your time to be wasted (this is a proverb we have coined at Fing). You will never be able to plan everything down to the last moment. Nevertheless, beyond applying experience and realism to the task of estimating timelines, meeting deadlines is a matter of communication: a seemingly tight deadline is a useful tool to make some of them take greater responsibility for their input.

Beyond these general considerations, the three facets of a Self Data experiment are all conducive to creating delays. The most fraught with risk is probably the data sharing aspect. Every organization has to carry out an internal, cross-departmental process voluntarily (personal data is a legally sensitive subject) — and it will rarely be their operational priority. You must find your strongest collaborators within each organization as early on as you can. Do not hesitate to consult Fing's work on data portability for support, especially our “[practical guide to data portability](#),” which outlines the process in detail.

A SELF DATA EXPERIMENT: 10 CARDINAL RULES TO RESPECT

01 ACHIEVE BALANCE BETWEEN DISPERSION AND CONCENTRATION, YOU WILL

Is choosing a theme before launching a Self Data dynamic a barrier or a lever to its development? The main advantage of not having a fixed theme pre-launch is that you can consider bringing a more expansive range of data holding organizations and partners on board. The experiment will thus be more focused on individual data management, and less on embodying “Self Data and X” (the energy transition, mobility, etc.). The disadvantage of this type of approach: dispersing your efforts. Tester and data densities are sometimes not coherent. Focusing on an explicit theme enables you to garner the participation of the right actors very quickly, to sit them around the table from the get go. Above all, a stated theme enables you to make Self Data more concrete for the user by associating a temporarily fluid technical solution to a real problem. Stating a theme creates space for something new to emerge — it is your territory for learning — which also has the benefit of increasing your competitiveness and sharpening the visibility of your efforts relative to other cities on similar Self Data paths.

02 KEEP THE FAITH IN SELF DATA, YOU WILL

Using Self Data to transform the digital economy directs means directing your efforts toward a very desirable and worthy future outcome. Keep that in mind as you grapple with the many obstacles that can stand in the way of your practical pursuits.

03 TREAT NEWCOMERS WITH CARE, YOU WILL

Take care of your communities: partners, reusers, and testers will form the bedrock of your experiment, throughout the months (and sometimes years) to come. You will need to patiently answer the same questions, address the same concerns, and continue to move forward without taking “one step forward, two steps back.” Anyone can thwart the experiment at some point if they are not sufficiently informed, especially outlier stakeholders who are following the project from afar or come on board later in the process. Consider delegating responsibility, training accomplices, and recruiting Self Data champions and ambassadors as part of your efforts to establish a more horizontal space for debate.

04 AVOID MISUNDERSTANDINGS, YOU WILL

Self Data is a complex concept. We have fond memories of one partner who thought that the Self Data experiment would grant them access to all the testers’ personal data, and a reuser who told us they couldn’t begin developing services because they had no data to work with . . . Avoiding misunderstandings will save you precious time. This is why you have a FAQ: to answer testers’ and partners’ questions upstream. Literacy around Self Data applies to everyone, even to the most senior data experts.

05 DOUBT THE MAGIC OF DATA, YOU WILL

Shared personal data should foster the emergence of innovative uses. But you will have to avoid the trap of underestimating the gap between what you can feasibly do with data, and magical thinking about “what we could do.” Some specific properties of data (type, frequency, latency, etc.) can block certain uses (such as data crossing, for example). Some uses imagined in workshops may be impossible to implement using the data available. Crossing datasets may remain a fantasy. Do not underestimate the time it takes to obtain data, nor the difficulty of mobilizing reusers.

06 THOU SHALT NOT LIE . . .

Once you have put together your final experiment scenario, it will seem ideal. You will have made the effort to please everyone, you will have hybridized several scenarios into one. Beware. If it’s too good to be true, you risk disappointing partners and testers. Avoid half-decisions: make explicit choices (about the target audience, the services developed, the means with which you will enable their co-design, etc.). Your experiment will simply be better.

07 . . . BUT THOU SHALT ALSO DARE TO DREAM.

An experiment Self Data must evince a considerable degree of ambition. After all, it’s a very new concept, you have to keep dreaming to embody it. And do not dream alone: imagine simple but entertaining uses with the testers and partners. You will be dedicating a lot of brain power to understanding Self Data, and lots of time testing it and implementing it: make that time fun.

08 ASSEMBLE A VARIETY OF DATA HOLDERS, YOU WILL

One thing you must decide immediately: can a data holder also lead the experiment? Local public actors, decidedly so, in our view. Localized public entities can bring together actors with diverse and often competing interests. (Fing has also been able to achieve this.) Putting a private organization in the driver’s seat can lead to problems enlisting allies. Some will not be able to join a “private” initiative (eg: led by a company whose stated mission is public service), some will refuse to join forces with their competitors, etc. Your experiment should include at least two to three data holders who agree to share the private data they hold with testers. But if you have only one holder who has a lot of data to share (for example a city), the experiment can still work if the foundations of the experiment are solid.

09 DOCUMENT REGULARLY, YOU WILL

Thoroughly documenting (and communicating about) each step in your experiment, and capitalizing on your learnings, will show others the way. Create common knowledge! Self Data has become more popular recently, you are not the only one to consider launching an initiative: cooperate, share your experiences, create synergies. In addition to carefully documenting your progress and status, do not forget to create a log detailing exactly the types of data you are using (fictitious datasets, data repository sources/geolocation data, etc.), including the means through which they were shared — you will be laying the groundwork needed for reusers to exploit the data successfully.

10 PUT TOGETHER A DREAM TEAM, YOU WILL

You will need a team whose members, together, possess a wide variety of skills. You will also need a dedicated coordinator working full time, of course, or even two — and they should be working for an organization that can communicate effectively with every stakeholder. These entities may be employed by the city, work for a DSI or in the field. Here may be the perfect opportunity to recruit a third party to lead the experiment! You will also need a “guru:” a highly placed person who can project the Self Data message loud and clear. You also need a killer tech supervisor, motivated researchers, expert designers, a facilitator to oversee the test cohorts, and another for the innovation clusters, etc.



**APPENDICES
READY-MADE
METHODOLOGIES**

SELF DATA SEMINAR: METHODOLOGY

Objectives :

- » The participants come away with a solid understanding of Self Data and its potential.
- » The participants are able to tie Self Data together with their specific local concerns.
- » The participants have the desire to pursue the topic of Self Data further, and consider data and data uses.
- » We have a list of specific challenges for Self Data to respond to;
- » We have drawn up the “guest list” for future workshops, when you can open the debate to a wider audience, conduct interviews, have meetings....

Participants :

- » 30

Time required :

- » one full day, 10am-4pm, with lunch break

[Download the seminar template here](#)

Morning (2.5 h): Plenary session; presenting and self-teaching the basics of Self Data

5min - Welcome and Icebreaker

Let them play the “5 min 20 data” game; or maybe try the easier “data in your wallet” game, which consists of asking participants to “Look in your wallet, find a card that you believe implies personal data in some way, and hold it up.” Examine the cards being brandished, choose a participant, and ask “What kind of personal data do you think this card implies?” Other participants help answer. Repeat the operation two or three times by choosing different cards (ex: a grocery store chain loyalty card will lead to consumption data (product names, quantities purchased, the number of people in the household, etc.) while a credit card will point to purchasing data (places, dates, amounts, etc. Drill down to the details as quickly as you can: it’s a great way to reveal how much personal data we generate offline, rather than online..

1h - Presentation of the Self Data concept; Q + A

Use our [content](#) to prepare your [slides](#), which will outline the general context, the stakes associated

with Self Data, examples/use cases, the value to be obtained by stakeholders, the international outlook, the GDPR and the right to data portability, etc. Get ready to answer a few typical questions about data monetization, ownership, revenue models, and more. The Self Data [FAQ](#) is a potent resource.

15min - Exercise 1 - Open question “What kind of value do you think personal data holds?”

You can circumvent the data resale issue by organizing a small exercise: after asking the participants the above, have them write their answers on a sticky note (one note = one idea). The facilitator then takes a few minutes to build a wall map by reading the notes and grouping them according to the answers. This is an opportunity to reveal how each actor sees personal data and correct preconceived notions about how individuals monetize them.

30min - Self Data Cities: implementing Self Data on a citywide scale

After defining Self Data, it is time to clarify the major issues for cities and why a regional approach can be fruitful. You can use our slides

and of course the first chapter of this booklet. Now is also the time to present the calendar of upcoming events and the dates of the next workshops.

10min - Exercise 2: “In pairs, describe what Self Data represents using any means you wish: a diagram, drawing, mime, a poem, a cloud of keywords, etc.”

This exercise gives you some time to catch your breath, and lets participants discuss what they have just heard amongst themselves. This ensures that you have been clear and that everyone has grasped the concept of Self Data. See what your participants come up with.

30min: A SWOT analysis – Smokescreen and knee-jerk objections to Self Data

A Self Data seminar means that participants will be questioning the notion of Self Data, and asking countless questions that you won’t necessarily have answers for, and that there will also be some controversy. The debate can get heated! To avoid posing as the sole “advocate” for self-data, and to spread the responsibility around, a SWOT exercise will involve everyone in imagining Self Data success.

Put participants into pairs, and then :

- » give everyone two minutes to list the risks and opportunities associated with Self Data;
- » take 15 minutes to allow them to exchange ideas with another pair;
- » stick their responses (on sticky notes) on a wall, making sure to group common ideas together.

To sum everything up, ask one volunteer to play the devil’s advocate, and another to play the Self Data defender. They can use the lunch break to group the identified risks and opportunities together by category (for communities, for businesses, for individual citizens).

Lunch (1h) : Break

Before you continue, give the two volunteers a chance to sum up the risks and opportunities for each category.

Afternoon (2.5h): Looking forward – “Challenges/Data/Uses” themed workshops (small groups)

Objectives:

- » Elucidate the issues and challenges specific to each region according to the theme (eg: mobility, sustainable energy consumption, social welfare, etc.);
- » Discuss existing regional initiatives;
- » Link all this to Self Data, in terms of (potential) use cases and data.

To make this part of the day pass more smoothly, before the event make a small list of issues related to the day’s theme that are likely to concern local public actors from each region. These challenges must correspond to the themes their projects will address and their strategies for 2020. For example: “increase soft mobility,” “facilitate intermodal passenger transport,” “calculate the carbon footprint of transportation options in our region,” etc.

I/ Issues/Challenges (plenary) - 30min

This part is essential. The themes that emerge can always be refined as you go forward, but they will inform your workshops and the experiment for

many months to come.

Ask participants the following question: “What kinds of (eg) mobility/transport challenges would you personally like to address in your region? If possible, express these problems using an action verb.” (When the theme is mobility, ideas might be things like “help seniors use public transport to go shopping;” “encourage young drivers and learners to opt for carpooling.”

Participants take 5 minutes, solo, to jot down one challenge per sticky note; then each participant presents their ideas, and the facilitator maps them out on the wall by theme.

You can then put the participants into groups, each working on a different challenge/issue.

For the following stages, use the template hyper-linked at the beginning of this appendix. Remember to modify it to fit your theme, and print out several copies on A1-sized paper.

II/ Link issues and the potential ways Self Data might address them - 30min

The groups begin by noting down their theme on the A1-sized sheet, and the issues associated with it, and then iterates ideas for projects and the people they can draw on as resources in their

region who are already working on these issues. After that, they discuss the use cases that could emerge to address each issue.

Prepare yourself by consulting our document detailing the 7 domains of data use, and ask participants the following question: “What kinds of services and use cases could you use to address this issue?”

For example:

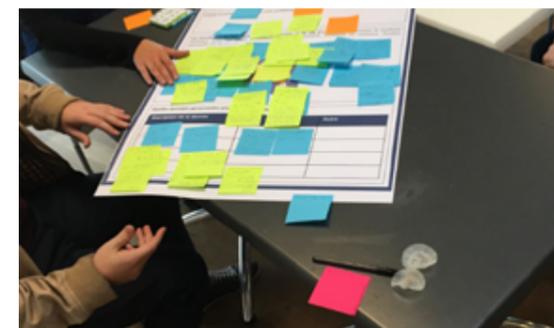
Contribution: “share data to allow individuals to complete transportation surveys;”

Management: “make life easier for citizens, mobility-wise;”

Awareness: “evaluate the carbon footprint of the region, and reduce it;”

Decision making & taking action: “help citizens manage their mobility budgets.”

No need to go further than this; you are not there to come up with specific use cases, you are there to generate the maximum number of ideas for services that could address the issues. If some of the use cases are not related to personal data, note them down . . . they may very well come in handy!



III/ Associate the Self Data possibilities to the data available in the region - 30min

Then take each use case, and try to reveal the data behind it. Tell participants, “The ideas you just had will necessarily need to draw on personal data – which data?”

Now it’s time to fill the mini spreadsheet on the templates; without going into too much detail, you will all be able to identify some data that would make sense to share with individuals. Do not focus solely on personal data – also look at data repositories, open data, etc. Not sure if a dataset exists? Note it all the same – you’re here to understand the possibilities within Self Data, not to produce content you will use right away.

IV/ Plenary session - 30min

The last half-hour of the day will be devoted to summaries delivered by each group, reminding participants of upcoming events, and holding a round table discussion to figure out who was missing that day and should join in the next workshop. Make the network work: ask participants to contact the people they think about, and give them a template email you have prepared to that effect.



DATABLITZ (DATA HUNTING) WORKSHOP: METHODOLOGY

Objectives:

Based on the “Wanted Data List” methodology produced as part of a crossover between Fing’s MesInfos and Infolabs projects, this workshop will produce a wish list of data that would be useful or that participants think has some kind of association with a particular topic. There will, of course, always be some focus on personal data.

Sometimes you will reveal data that has not even been computerized yet. There are actually two possible scenarios: either the data you identify already exist (you know where to find them), or you would love them to exist, but they do not yet.

Before the workshop, data holders can do some “homework” and look at what they have as data in their information systems on the chosen theme. Now is the time to invite CIOs and IS professionals for a coffee, and invite them to the workshop!

Participants:

» 30-40

Time:

» 3h, plus one 15min break

Stage 1 (20min): Introduction

Thanks to the first seminar, you have hopefully widened the circle of people involved in the experiment process. So you must explain the concept of Self Data again, briefly, to bring everyone up to speed. This will also allow everyone to refresh their memory.

Stage 2 (40min): Summary of prior workshop and further development

You will not be starting out with a blank slate here: thanks to the earlier seminar, you have a range of issues and challenges to iterate on during this workshop. After summarizing the coming year’s key focal points, take a few minutes to present the 3 to 4 challenges discussed in the previous seminar, which you will return to in a plenary session shortly after. Now is the time to reformulate the issues with the relevant actors.

Spend ten minutes per challenge maximum if you have four. If you created a summary of the first seminar’s takeaways, then you can present one slide per challenge/issue that includes notes about relevant actors and data, and the early use case ideas. The slides will enable participants to more readily join in the discussion, and offer you the opportunity to ask newcomers about their relevant contacts and ideas about existing initiatives that

feel should be included in future workshops.

Break (15min)

Before the break, tell participants that when they come back they will have to pick a challenge and sit down at that table (1 table = 1 challenge = 1 group). Take advantage of the break to place a template sheet for the groups to fill out on each table (see the figure in Stage 3); you can either prepare new templates on the spot on giant sticky notes, or print them out beforehand on A1-sized paper.

Stage 3 (1h45): Identifying relevant data

I. Data list (1h30)

Each group should ask themselves the main question: “What information will be needed to respond to this challenge?” Ask them to be as specific as possible: “We can start by listing general categories of information, but actual ideas about concrete data would be better (ex: an item like “private sport lesson fee information” is too broad; but data relating “length of commitment/number of lessons paid/number of lessons taken/number of lessons remaining” is better).

1) Begin with 5 minutes of individual reflection: each participant uses a sticky note to jot down What (name)? – and possibly Where (holder)? and How (access)?. They share their ideas with the group.

2) What information do we already know relates to this challenge? (eg: “education history”) – Where can I find information that would help me address this challenge?

3) What further information do I need to respond to this challenge? How can I locate that info? How could I obtain it?

» Do not hesitate to get online and see what kind of data might be available in personal profiles from service providers, for example

» Do not hesitate to tie in possible data repositories and sources of Open Data.

NB : Focus on the data. Potential uses are not important at this stage – what counts is to list the data and the potential holders that emerge when the challenge is placed in context. The facilitator picks up the sticky notes and notes down the details regarding the types and sources of data they contain. The best way to keep track of the relevant data sources and types is to have the groups use the template below that you will have placed on each of the tables.

Nom du défi / thématique :				
Post-its des participants	Information Qu'est-ce que j'ai besoin de savoir ?	Données Quelles données me permettent d'obtenir l'information ?	Détenteurs Qui détient la donnée ?	Accès Est-ce que l'accès à cette donnée est facile/moyen/difficile ? Vous connaissez quelqu'un qui sait ?
<input type="text"/>				

II/ Preliminary list of uses (15min)

The objective of this last part is to bridge the data types/sources with possible uses and potential services (you will be exploring this in the workshops that follow, so no need to go too far – this is about creating a vision).

Instructions (choose one): “Imagine a use for the data/What kind of service would enable us to address this challenge? What would it do?”

Ask participants to avoid giving one-line answers. Tell them to have fun! They should come up with :

» a fictitious name for the service

- » a one-line pitch/tagline that encapsulates the service concept
- » 3 key features.

Summing up (15min)

The participants return to the plenary seating arrangement and now one spokesperson per group takes a turn summing up their findings. Describing every single piece of data would take far too long. The spokespersons will therefore focus on 3 to 4 types of interesting data (the more original or surprising, the better) and the uses they have envisioned for them.

To generate value from the workshop proceedings, scan the results into a spreadsheet, and if possible make an easy to read map visualizing the results. Here are some examples.

> Visualizations based on data present in organizational information systems:

- » [the data visualization from the MesInfos pilot \(2016-2018\)](#)
- » [data visualization from our MesInfos experiment \(2013/2014\)](#)

Our most “prospective” visualizations:

- » [Self Data Cities \(energy, mobility, education\)](#)
- » [health](#)
- » [energy](#)



IMAGINE I WORKSHOP METHODOLOGY (AKA “SERVICE DESIGN WORKSHOP”)

This methodology was developed in partnership with the [BAM Collective](#).

Objectives:

- » Foster the emergence of a dozen service concepts, select a few, outline them, and create their usage scenarios.
- » These services must rely on personal data to generate value for individuals or groups of individuals. You will need to have completed the “DataBlitz” workshop beforehand (see previous methodology) and to have prepared either data visualizations or at least created lists that you can share easily.
- » Do not worry now about security, privacy protection, etc. The workshop is focused on generating uses. The coming workshop will be dedicated to sharing frameworks.

NB: If you have not compiled a data list previously, or created a data visualization, you will not be able to conduct this workshop. You will probably be able to imagine digital services, as many design studio methodologies do today, but the point of these “Imagine I” workshops is to construct Self Data services. Feel free to space out the interval between the Datablitz and this workshop to give you the time you need to produce the visualization and enrich it with data that might not have been recorded.

Participants:

- » 30-40

Time:

- » 3.5h, with one break

Les templates nécessaires à l’atelier sont téléchargeables [ici](#) (vous aurez à en modifier certains) :



- » [Service ideas template](#)
- » [Service template: detail](#)
- » [Persona template: Francesca](#)

Introduction and Ice-breaker - Plenary (20min)

- » Quick presentation of Self Data; information on the objectives of the workshop; reminders of the challenges that have emerged.
- » Each participant draws a “data” card inspired by the visualization produced after the Datablitz. The facilitator says, “Imagine that you can use or do something with this data, either by yourself or in collaboration with others. What would you like to do with it? You have 3min to think, then present your use idea in a few words.”

This is an opportunity to go around the table and trade ideas about the data that are familiar to Datablitz participants. If the participant cannot come up with a way to use the data, ask them to simply explain the thoughts that they associate with the data and how the data might potentially be combined with other data.

Stage 1: Challenges to services (40min) - group work

- » Reminder of the challenges and movement into groups (1 challenge = 1 group) - 10 minutes.

- » Generate service concepts - 30 min.

Materials: one A4-sized copy of the “service ideas template” per table

Have the groups begin by iterating on the challenge: What sub-challenges does it entail? What aspect of the challenge should we tackle? What is the problem inherent in this challenge? (education, support, etc.)

Ask participants the following questions:

- » “We have identified several significant challenges experienced by citizens in your regions; now, what concrete services can you imagine that would enable individuals to contend with these challenges?”
- » “Can you formulate these service ideas like this: ‘It is a service that would do X . . . but also Y, and Z, and that we could also use for/like A.’”
- » “What form would that service take? (app, site, object, social network, etc.)”

Participants may think individually for one or two minutes, then use sticky notes (1 post-it = an idea) to share ideas with the group. The facilitator completes the A4 “service ideas” template. Each group can imagine 1, 2, 3, 4 ... ser-

vices per challenge, with the facilitator’s help.

This stage is fairly short, and it should go fast – you want to generate the maximum number of simple ideas possible in a limited time without questioning each one of them. Sorting will come after.

Stage 2: Services and Data (10min) - Plenary

You present the data visualization you created after the Datablitz. Get it up on a screen, and distribute the list of data you compiled (as a visualization, as a simple list, as a data map, etc.). Remember to print out enough copies (at least one per table).



An example of a visualization created for the workshop, with printable “data cards.”

“You have probably come up with a lot of interesting service ideas. Some are probably achievable now, and some will probably take more time before they see the light of day ...”

... and so, and for the rest of the workshop, we are going to compare your ideas with reality; to work, services destined for use by individuals have to be driven by data. Thanks to the Datablitz workshop, we have identified and documented a number of these data, the data from the “personal [NAME OF YOUR THEME] data landscape.”

Stage 3: Services definition and development (40min) - group

Back in their groups, participants choose an idea they can all agree to work on. Remind them to select an option that will require the use of personal data.

Have the groups begin to refine their service concept. Ask them to define and develop :

- » a pitch,
- » the kinds of data the service will require,
- » the service delivery mode (app/platform/object, etc.),
- » the target audience,
- » the benefits it will afford to users.

To organize your thoughts, use the first two pages of the service template: detail.

Break (15 min)

Stage 4: Create use cases (1.25h) - group

I/ Participants choose a service group/select a persona (15min)

If you think the participants are motivated enough, invite them to switch to another table/join a different group. A short presentation of the former group's imagined service follows. “We invite you now to either remain where you are seated, or to join a different table where a different concept was imagined. At least one participant must remain at their original table, to present the concept. Wherever you decide to sit, the objective of the next phase will be to script the service, meaning that you must be able not only to describe its features in detail, but also some specific, concrete use cases/scenarios for it.”

If the group composition has changed (this is a time thing: keep the same groups if re-summarizing the service concepts isn't feasible, but switch things around if possible) and allow participants to familiarize themselves with the services at their new table.

1) What does the service do?

- » Do we agree with the principles the service is based on – do we think it will create value for the individuals who use it? Are we satisfied with the features it offers? Is something clearly missing?
- » If we agree on the above, we continue to work on the service as described. If not, what are the service's most relevant features? Which audience(s) does it target? Do we agree with the target audiences described?

2) Persona selection

At the end of this stage, participants will select 1 or 2 personas (from the persona cards available, or from their imaginations). the personas will enable them to create a concrete picture of how the service might be used, by whom, under which circumstances, to what benefit.

You can use the “persona template: Francesca” that you will have adapted to suit the day's theme.

II/ Deepen the persona profile (20min)

Describe the characteristics of the persona (or personas) you have chosen:

- » Who are they? (age; civil and professional

status; relationship to the theme, etc.)

- » How do they relate to the theme? How do they experience it?
- » Which needs (or motivations) would compel them to use this service?
- » What obstacles will the service you are offering them need to overcome?
- » What are the service's principal flaws?

III/ Use a persona to describe the use case/service scenario (40min)

“Plot the story of your persona using your service on the usage scenario timeline. You can tell the user journey in 4-5 steps (what does he see on his screen, etc.). Use these questions to help you:

- » “Which situations are problematic for this persona today? Explain the challenges that your persona is dealing with at the moment (within the context of the service).”
- » The purpose of this step is for participants to discern contexts for service use that will benefit the persona and are relevant to the persona's concerns. “In each of these situa-

tions, explain how the service is used and how it addresses the persona's concerns: how will the persona use it? How does one access the service (via which medium: mobile app, online platform, etc.)? What stages does a user go through when using the service?”

- » “Back to the data: to do all this, what data will the service need? Which information will the persona have to enter manually, which data will be collected automatically, etc. ... and when?”

“You should now fill in the last page of the Service template - detail (if you have people with artistic abilities in your group, you may ask them to design some of the ‘screens’ that users would see).

Step 5: Summing up/Services Pitch (10min) - Plenary

Each group shares the results of their work as briefly as possible: name of the service, pitch, some features, 2/3 items of data required for it to function, persona name and 2/3 key steps the persona needs to follow to use the service properly.



THE IMAGINE II METHODOLOGY (AKA “WHICH MODEL OF GOVERNANCE FOR WHICH USE?”)

Objectives:
Take a more global approach to the services and use cases previously imagined by assigning a model of data governance to each one.

- 1) Become familiar with the existing options.
- 2) Consider the “back office” operations the use case will involve, and create a governance model to scaffold it.
- 3) Make sure the model overcomes at least one obstacle.

Participants:

» 30-40

Time:

» 3h.

[Download the workshop template and cards here.](#)

Stage 1: Setting the Scene (1h) - Plenary

» 10min: presentation of the Self Data Cities project and workshop objectives

» 5min “Imagined services:” spend a few minutes presenting the use cases created at the previous workshop. You will be delivering a brief summary of each, but you must have a detailed summary sheet for each service printed out for later in the workshop.

“At the previous workshop, we imagined N service concepts to meet the challenges. We will paint you a brief picture of each one (name + tagline + main features) so that you can associate these use cases with one or more governance models.”

» 5min “Why the current situation cannot continue:” This is a tricky foresight exercise, because it asks participants to project their minds forward, and imagine governance models that don’t necessarily exist today, but could if the actors so desired. So you begin the foresight process by explaining why it is important to discuss models today . . . by revealing what might happen tomorrow. Examples of disaster scenarios:

- The “Money, Money, Money” scenario: Everyone starts to want to monetize their

data, for real. There is a demand for data, a service that facilitates their resale, and clients lining up to buy them. Companies become individuals’ customers, and individuals sell their data. Everything is monetized, everyone thinks in terms of data ownership. All the while, the gap is widening between those who have the choice to maintain their privacy, and those who need money and sell their data to the highest bidder.

- The “A package of PIMS (per person!)” Or “The doorway to my data” Scenario: The notion of Self-Data is propagating and enjoying a certain measure of success . . . people have the choice between many data sharing solutions, and tend to open multiple accounts across the platforms available (yesterday multiple devices, tomorrow multiple personal clouds!). Companies, like insurer Maif (an investor in CozyCloud) promote one platform over another (EDF uses X, Société Générale uses Y); each platform allows individuals to access some of their data according to its partnerships (I access my data EDF and Maif data on CozyCloud, but not my Société Générale data), but few allow access to all of my data. To really get an

overall picture of where my data is and who is using them (the famous 360° perspective) people have to open multiple accounts on multiple platforms, and self-service application markets remain fragmented. This is a data bazaar: Self Data exists, but takes up a lot of everyone’s time and effort . . .

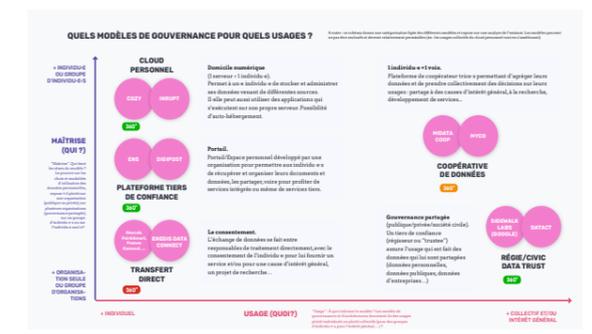
- The “Facebook, savior of Self Data” Scenario: Facebook proposes a PIMS, and everyone loves it. Millions of users are given a personal cloud backed by unlimited resources to throw at bugs, design, and endless uses! The only downside: Facebook’s business model has not changed, and it’s making a fortune from data resale. By virtue of his complete appropriation of a concept dreamed up by an obscure, little, long-forgotten French organization, Mark Zuckerberg becomes the standard-bearer of Self Data.

» 40min “Which governance models exist?:” Presentation of existing models of governance sheets (data co-operative, civic data trust, etc.).

“How to avoid that kind of extreme scenario? It’s time to put together new governance

frameworks for our personal data. We can base ourselves on a few ‘off the shelf,’ existing models, but as each has its shortcomings and advantages, the idea today is to draw from a few different data governance scenarios, embody them via the use cases that we developed during the previous workshop, and create alternative structures that are even better!”

For an “off the shelf” definition of each governance model, please consult Chapter 1 of this booklet.



Stage 2: Building an Alternative Governance Model (1h10) - groups

One group = 1 model + 1 use case.

Each group is allocated a use case. Before this workshop, you will have determined for yourself which models seem the most appropriate for each (collective uses for data imply data co-operatives, trusts, etc.).

“You have a service (ex: Toque Verte) and a data governance model (eg: the personal cloud). Concentrate today on the model of governance: the idea is to improve it, to modify it, to make it a desirable model that you would like to see emerge in your region to help you enable individuals to manage their data effectively.”

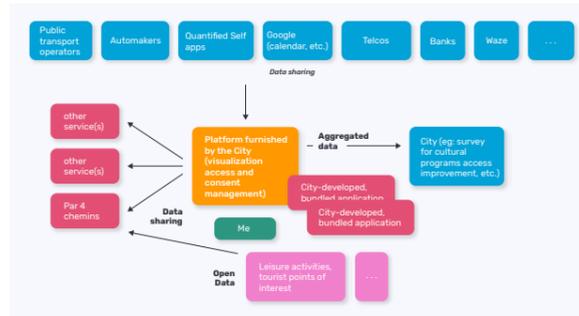
“Abstract reflection on a particular model makes no sense: you have to ask yourself ‘What does it do?’ This is where the use case helps you. Use it to guide your understanding of the model’s purpose, who it serves, etc. . . . so you are not just explaining some structure, you are illustrating how it can support uses that the inhabitants in your region will find useful.”

You can use the first two pages of the [template here](#).

“Use the template to answer this question ‘Why is

this model relevant to our use case? What does it offer?’ You can then take time to put together a new model, even a hybrid.” (This is the time to distribute the sheets explaining the ‘off the shelf’ models that exist. You can find visualizations of each in Chapter 1.)

“Describe where the data is stored, to whom it is shared, in what ways. You can represent it as a diagram in the dedicated part of the template by using the icons on the next page. Qualify as precisely as possible who the model’s key actors are. You should come up with a complete schema, like this one describing a service called ‘Roundabout’ that uses the ‘trusted third party’ platform model to manage data sharing with individuals.



If you have the time, consider describing what the model is transforming from the regional actor’s perspective: “are we making an actor obsolete?”; “This actor brings new skills,” etc.

Stage 3: Putting the Model to the Test – How Can it be Implemented? (30min) - group

» Facilitators choose from a list of ready-made obstacles that best fit the alternative model participants have designed. You can rely on the visualizations available in the template: data aspirations; destruction of data centers; generalized monopoly; rebellion. Do not hesitate to create others.

“How does this obstacle affect your model?” This pure foresight exercise allows participants to discern legal safeguards for their model, imagine the controversies it may create, and consider some initial responses to them.

» After thinking about the obstacles and describing a model that has been able to overcome them, take a few minutes to study the pathways opening towards making that model a reality. What milestones, what levers would be needed for that model to emerge today?

Stage 4: Summing up (20min) - Plenary

Each group summarizes their use case and the model that emerged, including the obstacle it overcomes, and the steps to making it a reality.



**EXPERIMENTAL SCENARIO
METHODOLOGY (AKA
“IMPLEMENTING SELF DATA IN
YOUR CITY”)****Objectives:**

- » Describe Self Data & “theme” use cases that can be developed via experimentation.
- » Converge the interests of (potential) experiment partners with those of the community and the city’s inhabitants.
- » Understand Self Data experimental scenarios.
- » Understand the parameters (budget, actors, partners, etc.) and the timelines relative to these experiments.

Participants:

- » 30

Time:

- » 2.5h

[Download the workshop template and cards here.](#)

Stage 1: Introduction (20 min) - Plenary

A little overview of what was produced during the past workshops will get everyone on the same page. Cover the concepts of Self Data, Challenges, Data, Use Cases and Governance Models.

Be clear about the stakes attached to this present workshop: it is a key moment for stakeholders to use all the work carried out previously to inspire their creation of one or two experiment scenarios that enable their city to move towards futures that are desirable, and also feasible.

“A Self Data experiment has three fundamental requisite components:

- » data holders who agree to share data with the individuals they concern,
- » testers who will be able to reuse those data;
- » tools to enable testers to reuse their data (platforms, third-party services, etc.).

“This workshop enables you to precisely define different types of experiments: number of testers, timelines, technical solutions, data requirements, budgets, etc. It will also serve as a means of clarifying the stance of any potential experiment partners: are they ready to join you? Under which conditions?”

Stage 2: What We Want to Learn/What We Want to Avoid (30 minutes) - Plenary

To prepare for the workshop, you will have to spend some time developing some axioms for each experiment. After several months of iterating and discussing the experiments in workshops and beyond, the objectives that each experiment seeks to achieve and the conditions that must be established to guarantee their success ought to be fairly clear. Get some help from colleagues to iterate these axioms yourself before the workshop.

Start by introducing them to the participants. For example, “The experiment in La Rochelle must ...

- » meet two objectives:
 - test the validity of the Self Data as a means of controlling personal data;
 - explore the universality of Self Data (beyond testers – digital mediation);
- » involve:
 - about fifty testers who are employed by the city, the Post Office, or Enedis;
- » have testers test uses cases related to mobility:
 - My Mobility Budget; Shared Mobilities; CO2 Coach.

Then, compare your list of axioms to the participants’ using the following exercise. The goal is to find ways to converge the interests of (potential) partners in the experiment with those of the community and the inhabitants.

Answer these three questions:

- 1) What would you like to learn and/or create and/or demonstrate via a Self Data experiment on [theme] in [City]?
- 2) What do you need to give the go ahead for an experiment like that? (ex: legally substantiated document detailing where, how, and when the data will be shared)
- 3) What do you want to avoid at all costs? (ex: an experiment that only starts in 3 years, not being able to recruit enough testers, etc.).

Everyone silently prepares their answers (use different color sticky notes: 3 colors, 1 color per question) - then shares them so that the facilitator can group them on a wall.

Stage 3: Producing Experimental Scenarios (1h30) - Group

You can then form two to three groups. Each

group will work on an experiment scenario that has a different starting point.

In the case of Greater Lyon: “N testers each have a personal cloud provided by the city. They can download “X” service from the personal cloud store, which has been developed by the Greater Lyon.”

Each group should have an experimentation scenario template to scaffold their efforts. Also consider printing out the summary sheets of the use cases that have emerged.

“Your scenario building exercise will be to ask and answer questions related to ...

...data (who are the 2 to 5 holders the experiment absolutely cannot do without? Which data would they share? Which API do they use?)

...testers (who, how many, geographical perimeter, recruitment/facilitation modes, user feedback)

...services and uses (Which services will be implemented and tested during the experiment – NB: use the hypothetical use cases and models of data transfer and governance you learned about previously – and how, eg: development from A to Z, enrich-

ment of existing services, competitions, etc.?)

...actors (Who are the data holders? researchers? funders?)

...and timeline/budget (keeping in mind that the experiment will likely take place across different fronts: round table composition, recruitment, facilitation/research, services and development). You can take inspiration from the 'MesInfos Pilot' booklet to help you flesh out your ideas.”

Stage 4: Collective Feedback (10 min) - Plenary

Each group shares its experimental scenario in plenary. After the workshop, a lot of work will be needed to extract the material produced and reformulate it into specific scenarios. You will need to be able to produce an internally and externally shareable document ([see an example here](#)). This will be the basis for any agreement with the possible partners on one of the scenarios. The document can be amended and then presented publicly (perhaps in the form of a “Partnership Orientation File”) ... and then you can get the experiment started!



Self Data cities, the playbook

**"WHAT IF CITIES TOOK A CENTRAL ROLE IN RETURNING CITIZENS'
PERSONAL DATA TO THEM?"**

2020

mesinfos.fing.org

