SMIT STUDIES MEDIA INNOVATION TECHNOLOGY

Open Banking data in Smart Cities

Round table report

February 2021

VUB CHAIR DATA PROTECTION ON THE GROUND

in partnership with BNP PARIBAS FORTIS

SMART CITIES CHAIR

VUB joyn Belfius proximus Befimmo

ABOUT THE CHAIR ON DATA PROTECTION ON THE GROUND

The VUB Chair "Data Protection On the Ground" (DPOG) promotes the investigation into actual practices of data privacy in organizations and the dissemination of best practices. The focus of its research is on developments in smart cities, and the health, media, and financial sectors. To this end, the Chair compares practices in public sector organizations with those in the private sector, and organizations experienced in personal data protection with organizations that are making their first steps. In lectures, workshops, roundtables and other events, the Chair brings experts and practitioners together to stimulate the discussion of best practices.

The Chair is coordinated by the research center imec-SMIT (Studies on Media, Innovation & Technology) in collaboration with the research group LSTS (Law, Science Technology & Society), and supported by BNP Paribas Fortis. For more information, please visit the Chair's website at www.dataprotectionontheground.be.

ABOUT THE SMART CITIES CHAIR

The chair on Smart Cities is organized by the Vrije Universiteit Brussel, the Faculty of Economic and Social Sciences and Solvay Business School, and imec-SMIT (Studies in Media, Innovation and Technology), together with its partners Joyn, Belfius, Befimmo, and Proximus.

The main goal of the chair is to increase and deliver knowledge exchange between experts, public and private stakeholders on how to involve and incentivize citizens and other communities, how to work with data in a privacy-friendly way and how to ensure compelling use cases that have real impact on social and economic processes in areas such as mobility, retail and so on. More information on https://www.smartcitychair.be/ .

ABOUT SMIT

The imec-SMIT-VUB research group was founded in 1990 and conducts fundamental, applied and contractual research on IT, media and policy. Our focus is on research related to innovation, policy and socio-economical challenges. To this aim, imec-SMIT-VUB conducts user research, policy research and business analysis, making use of both qualitative and quantitative methodologies. The research group consists of two research programs ('Media & Society' and 'Data & Society'), which both are subdivided in three specific units. For more information visit https://smit.vub.ac.be/ .

ABOUT THIS REPORT

This report describes the results of a roundtable session in January, that brought 14 experts from different stakeholder groups together to discuss the use of open data in smart cities. The workshop was chaired by a representative from Mastercard (Helena Koning) and imec-SMIT (Prof. Dr. Jo Pierson) Under the Chatham House Rule, participants were free to bring in any topics related to smart cities and open banking that they saw fit.

The analysis for this report was conducted by Ruben D'Hauwers and Ine van Zeeland, with support from Koen Borghys; researchers at imec-SMIT, Vrije Universiteit Brussel.

REPRODUCTION

AUTHORS

Ine van Zeeland, ine.van.zeeland@vub.be
Ruben D'Hauwers, ruben.dhauwers@vub.be

imec-SMIT, Vrije Universiteit Brussel, February 2021

Contents

Introduction

This report describes the results of a roundtable session on 19 January 2021 that brought together 14 experts from banks, academia, government authorities, fintech and related industries. The main topic of discussion was the use of open banking data in smart city projects. The experts discussed use cases, needs and limitations, and trust issues. Divided into two groups (one discussing in Dutch and one discussing in English), participants were free to bring in any input related to these themes as they saw fit.

The following guiding questions were used to spark the discussions:

- Group 1: Purposes and means of open banking data in smart cities (in Dutch)
    - What are the needs of cities and governments? For which purposes would that be possible? What are convincing use cases?
    - What are challenges to implement open banking in smart cities?
    - What are possible solutions and conditions?

- Group 2: Trust in the sharing of financial data with the public sector (in English)
    - How can citizens be informed about the use of their financial data?
    - Is there sufficient trust within the ecosystem to share these data with the public sector?

Purposes of open banking data in smart cities

Needs and use cases

The needs and potential use cases of open banking in smart cities can be divided into different types:

1. Facilitating administrative processes

The sharing of data can **make the life easier of citizens, while improving processes of the government.** Some examples could be the following:

- The bank data on the income of citizens could be utilized to facilitate the subscription to social support and/or to tax details. This could make the process easier for citizens.
- The banking data could be used to make the controls of income details in tax declarations easier. The banks could provide input in the administrative process of tax declarations.

2. Supporting economic policies

Banking data could support governments to f**ormulate policies for the economy**, in different sectors such as local retail. This could especially be a driving force in the recovery after the economic impact resulting of COVID-19.

The data could be used to analyze data which can be of support in creating economic policies:

- The offering of online shops versus online shopping. The percentage of payments of consumers in the e-commerce sector and/or to foreign companies.
- To investigate the geographical origin of consumers, and for which type of purchases consumers visit a city.

- To verify models which make predictions on the type of visitors and transactions in supermarkets, in clothing shops, etc.
- To measure the economic activity in a specific region, by estimating e.g., what profile of people work in a certain region and what their purchasing behavior in a specific area is. An example can be the study of the impact of the harbor of Antwerp on different Belgian cities, and what the impact is on the taxes.
- To measure the impact of local events on the income of local merchants
- To measure the transaction trends of citizens in a specific area and period. This could be utilized to attract investors to specific locations or be utilized as an input for merchant platforms
- Rewarding local buying for consumers

3. Supporting sustainability policies

The banking data could be utilized to **support sustainable development**, both on an ecological as a economical level. This could be done by:

- Analyzing mobility with open banking data, to measure the impact of e.g., mobility policies.
- To provide incentives or reductions to incentivize sustainable behavior, such as utilizing the bus and/or to facilitate policies to reduce parking in the city environment.
- The payments of rent could be coupled to data on energy consumption. This way, the government could identify energy efficient buildings.

Challenges

An important challenge for sharing banking data is to ensure that the rights to **privacy and personal data protection of citizen**s are respected. There are different privacy enhancing technologies in the market currently that can support ways to ask for consent and anonymization. These still have challenges in ensuring the privacy rights are respected:

- Providing consent is only applicable in a certain time, and the willingness of citizens might change over time.
- Anonymization of citizens is performed, but a challenge related to this is that users can still be reidentified. The data of open banking is gathered by banks and private companies. They require the consent of citizens to re-use data.

An important question is whether the **combination of data** in a data lake of transaction information is desired. Major challenges occur in creating a combination of data:

- Different limitations exist as the consent and **privacy protection** of citizens need to be ensured, which reduces the possibilities to combine data.
- Special caution needs to be paid to ensure **re-identification** of citizens is avoided when data streams are combined.
- A lack of **standardization** of banking data is a challenge which is causing difficulties to interpret and to combine data from different providers
- A lack of **trust between companies** can also be observed, as competing companies might be scared that the data might be sold to their competitors.

The trust and knowledge of the citizens are crucial in this aspect:

- The **knowledge and understanding of citizens** about the sharing of banking data is limited, as there is not always an understanding on which information it concerns and to which extent the data is anonymized. According to representatives of the banking sector, the banking sector is very restrictive, and aggregation of data ensures the threat is rather low. Yet, the citizens do

not have an understanding of what this means Therefore, the explainability and intuitiveness of the data use is limited.
- The latter leads to a **limited trust of citizens in personal data sharing by the banking sector**. The trust differs from citizen groups, as older citizens might be more cautious to share their data compared to a younger generation. Additionally,

The **PSD2** regulation puts the citizens in the middle and enables companies to collect third party data when the consumer consents with this. The consumer could share data when, for example, given he or she could receive i**ncentives in return for sharing data**. This could provide opportunities, but the PSD2 legalities also has challenges as:

- PSD2 is **restricted only to payment data on the current account of a citizen,** meaning that e.g., payments with credit cards are not included. Thus, the data that can be opened through the PSD2 legalities does not provide a complete view.
- The participants in the round table added that a level of **reciprocity** is required from other sectors, to ensure that data can be shared among different players in the ecosystem.

Solutions

Digitization is an important driver to ensure the possibilities of data sharing. **Privacy enhancing technologies** play an important role to overcome the shortcomings of consent and anonymization.

- **Encryption**, which enables sharing data without disclosing the identity of the person.
- Share the data through **APIs** or through the sharing of only the analysis and not the raw data.
- A **posteriori feedback through a timeline in a data dashboard**. People could provide consent in the beginning of an interaction, but through the timeline they could be informed about the usage of their data in a later time, and the user could provide feedback on whether they would continue to allow these transactions. Thus, users would receive a notification “your data was used by bank x”, and they would be able to allow that this was performed.
- Another technological solution is SOLID, which ensures the **vertical disintegration of data storage**, where people could have their own data pod on which they could give access to different actors. Thus, the user could have the **control over their own data**. The challenge in making this technology widely available is to have sufficient incentives for the user so they will be able to use the data sufficiently. An equilibrium between privacy and convenience will be required.

Also, non-technical solutions can provide a way forward, by engaging citizens and by increasing the interests and trust of users to share their data:

- The willingness to share data depends to a large extent on the **goal for which the data will be used**. People might be more willing to share their data if they knew it would e.g. support cancer research.
- Additionally, **rewards or reductions** could be provided to incentivize the sharing of data. Additionally, a clear communication which is understandable for citizens is crucial.
- When data would be made available for the public use as **open data**, the willingness to share data might increase. This way, the benefits of the insights of the data could be returned to the users. Additionally, sharing the data as open data would ensure that users would become accustomed to the data.
- A clear **communication on the use of data** and how privacy of data is ensured is of high importance.
- **Independent control** on the use of data, for instance by a national supervisory authority, should verify whether the communication is correct.

Policies also play an important role in creating a solution for the sharing of data.

- An example is the Digital Services Act of the European Commission, which aims to protect consumers and their fundamental rights online, establishes a powerful transparency and a clear accountability framework for online platforms and fosters innovation, growth and competitiveness within the single market. Platforms will need to make their data available for free, in order to reduce the lock-in of data by monopolies.
- **Governments can stimulate the sharing of data** through collaborations with the financial sector and the local governments.

Trust and informing the citizen

The main questions in the second group were:

- How can citizens be informed about the use of their financial data?
- Is there sufficient trust within the ecosystem to share these data with the public sector?

At the beginning of the roundtable the concept of 'open banking' had been quickly introduced by a moderator as 'sharing banking data with non-banking partners', but one of the experts in this group felt a clarification was needed: "Open banking is about more than data sharing. It is setting up **an ecosystem** with different participants, where the total sum of the collaboration is higher than the sum of individual contributions. The value has to be clearly perceived by each participant." This expert and others emphasized that the ecosystem only works if everyone gets something out of it; a bank would not share data of its clients if there was no benefit to it.

From the perspective of bank clients/citizens, this means that sharing data should also be useful to them. But how does one make clear what the **added value** of open banking is to clients or citizens? They 'own' the data, but how do we make clear to them that they can do with it?

Communication about open banking

There should be no shame in being transparent about what happens with people's personal data. In that sense, it should be possible to be open about all the elements that go into open banking. The purpose of use should be clear as well. In the case of using bank data for public purposes, those **purposes should be beneficial** to citizens (so the data should not, for instance, be used to punish them for travelling during a lockdown).

Several experts brought to the table that **privacy notices** as they are now, are not effective in attaining transparency. People ignore them or find them too difficult to read, or to put it mildly: "Not everybody reads all the lines of a privacy policy." Official government publications about new uses of personal data are also not widely read.

To improve communication, '**normal language**' is needed and the added value of sharing data should be made clear **in a transparent manner**.

In some cases, people find out after the fact that personal data they had provided for one purpose is reused for other purposes they were not aware of, even if these were communicated to them in some form or other. In those cases, they may be "reminded that all data that are registered can be used against them." These types of **unpleasant surprises** may undermine trust and should therefore be prevented - another reason to be transparent about benefits and risks of sharing personal data. But even so, one can warn about potential risks concerning current uses, but it is difficult to predict for

anyone what possible future uses there may be that are currently not yet technologically possible.

The interplay between PSD2 and smart city projects

What about the other end of the equation: the smart cities, or local authorities, who could use the data from banks? An expert pointed out that they are not using the term 'smart cities': "Instead of 'smart cities', we talk about 'connected cities'. It is more inclusive, so we don't have to discuss what is smart and what not."

Terminology aside, an important thing to remember when it comes to the public sector, is that, according to the GDPR, the government is **not allowed** to work with informed consent, because they have other ways to obtain data from citizens, for instance by creating legislation or regulations. The government does not have to ask for consent to process personal data, e.g., not for tax purposes. If, for instance, social security data, health data, or national registry data are used for other purposes than what they were collected for, the original steward of the data (the authority that controls the data) needs to approve that new use, which is **a very strict process**.

There is some intricate interplay with the second Payment Services Directive (PSD2) here, because **PSD2 requires consent** from bank clients for the use of banking data, which public authorities are not allowed to use as a legal basis for processing personal data. However, 'smart' or connected city projects are often public-private partnerships. Non-governmental partners can still rely on PSD2 to obtain data from bank clients and process those data (perhaps in aggregated form) within a smart city project.

Even without having asked consent, there is still obligation for public authorities to inform citizens about the use of personal data. For most people, it may be more important to be informed and have some control over what happens with the data, to empower them. In that context it was mentioned again that Flanders is working on a project with decentralized identities and personal data 'pods', based on SOLID principles, in which people have more insight and control over how personal data is and can be used. Giving control and overview also adds to trust.

Socially acceptable use, trust and reputation

Elaborating on the Flemish project with decentralized identities and personal data stores, an expert explained: "For financial data, an important use case is to have more data about people's incomes. Now, government data on incomes are always two years behind and people's lives may have changed dramatically in the meantime, affecting their incomes." Having up-to-date income data would improve the relevance and proactive response of the government to people's financial circumstances. If this is not done with anonymized or aggregated data, but with data that allows identification of an individual, this can only work with informed consent. "If people know about the opportunities that are missed now, they will consent a lot more."

This illustrates that to convince citizens of the added value of sharing banking data with local authorities, the use of financial data should be **relatable** and have **daily relevance** for the people whose data are shared. The use should be **socially acceptable**.[1]

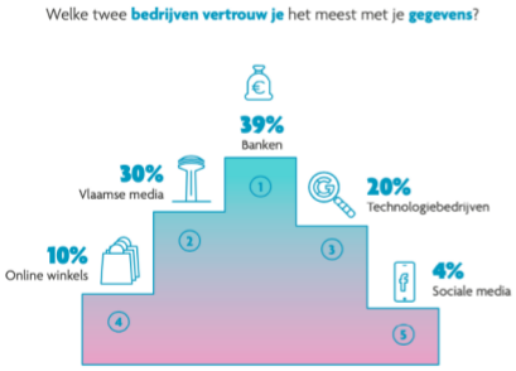
[1] In academic literature, the notion that certain practices need to be socially acceptable to succeed, is known as 'social license'. See, for instance: Aitken, M., Toreini, E., Carmichael, P., Coopamootoo, K., Elliott, K., & van Moorsel, A. (2020). Establishing a social licence for Financial Technology: Reflections on the role of the private sector in pursuing ethical data practices. *Big Data & Society*, 7(1), 2053951720908892. https://doi.org/10.1177/2053951720908892

Related to the idea of socially acceptable use, an expert stated: “Financial institutions are willing to share data, but in **a secured process** ensuring the respect of the privacy for the clients and managing **reputational risk**.” There are different facets to reputational risk. A partner in the ecosystem may use personal data provided by a bank in an inappropriate manner that makes newspaper headlines, implicating also the bank. Even if the partner’s use of the data is not strictly inappropriate, the wider public may feel uncomfortable with it, and implicate all ecosystem partners in their displeasure. Some partners may not even want to collaborate with others whose reputation are already tarnished. “The reputation of an ecosystem is as strong as its weakest link.” Another expert adds: “Reputational risk is difficult to quantify but we need to be careful with it. It takes years to build a reputation and seconds for it to be destroyed.”

In addition, government authorities can obtain access to data by law, as we have seen above, that not only the people but also the companies who collected the data might not agree to sharing. How to deal with that reputational risk is still an open question. Education can help raise awareness of the general fact that any data that are collected about you can be used in unforeseen ways.

On the other hand, the reputation of an ecosystem may also be built on the good reputation of the partners in it. Which institutions do people trust? A result from the 2020 Digimeter study:

Digimeter 2020 (Flanders – imec)



BETROUWBAARHEID VAN BEDRIJVEN

	16-24	25-34	35-44	45-54	55-64	65+	Totaal
Banken	41%	38%	37%	37%	41%	38%	39%
Vlaamse media merken	21%	31%	31%	36%	30%	32%	30%
Technologie bedrijven	30%	18%	22%	18%	18%	16%	20%
Online winkels	19%	12%	11%	8%	9%	5%	10%
Sociale media	8%	4%	5%	2%	3%	4%	4%

Welke twee bedrijven vertrouw je het meest met je gegevens?
Deze categorisering zijn samenvoegingen van twee of meer bedrijven

Apparently, banks are well trusted with personal data. An expert from a bank remarked: “This all depends on how the question is asked. Data leakage is not very common in banks, so yes, clients trust that banks store the data. But if you ask them if the bank can use their data for other purposes, like personalized services, that is not popular at all.”

Coming back to purposes, helping people to use their data for socially beneficial purposes can build bridges over the trust gap: once people see that their data about electricity use, for example, can be used to improve sustainable building plans, this will positively affect the discourse on contributing payment data. ‘Engagement tools’ can be used to provide transparency and give people the idea that they are part of the conversation. The personal data pods mentioned above can be such engagement tools.

The limitations of consent

Even when consent is an acceptable legal basis for processing personal data, the experts in the roundtable discussion expressed some reservations about the consent mechanism. Consent may become less useful when the **alternative in practice** is denial of common services: "We are evolving to a society in which you have to give personal data to get something from a business or government."

Another emphasized that "with consent, **responsibility is shifted** too much to the individual." Not all individuals will be able to make well-balanced decisions on what they are consenting to, for pragmatic, intellectual, digital skills, or power-balance reasons. Instead, we should be looking at other ways to accord accountability for what happens with personal data. We also need to pay attention to data minimization and publishing Data Protection Impact Assessments so they can be scrutinized.

Several experts noted a growing interest in society for accountability; the idea that those who use the data need to explain what happens with it, safeguard it and protect the rights and freedoms of those whose data are used. "These requirements are not only in the GDPR, but we also see them come back in other laws, like PSD2."

Procedures should be put in place to **improve accountability** and, in that way, social acceptance. "What if a bank provides data and the authorities use it incorrectly? There should be some **independent oversight or recourse** for individuals." Next to (legal) remedies, it would also help to have **standards for acceptable uses** and handling of banking data. In these ways, citizens can trust the system, rather than having to trust the players in it.

Conclusions and recommendations on communication and trust

1. Informing citizens

- The added value of sharing personal data should be clear to all participants in the ecosystem, including bank clients/citizens.
- Transparency should include informing citizens about the elements of the ecosystem and the purposes, as well as potential risks of sharing data.
- Current privacy policies and government publications are not effective in attaining transparency. They need to be improved or replaced by 'engagement tools'.
- 'Normal language' is needed.
- Consideration must be given to 'vulnerable groups', who may need more explanation.
- There is a role for education to raise awareness of the fact that all data collected about you can be used in ways you may not have imagined at the time of sharing.
- Empowerment for citizens means informing them and giving them some control.
- Communication is key to trust and social license. Engage with the public in dialogue and be prepared to counter negative comments.

2. Securing trust

- The use of the personal data should benefit the individual whose data are shared, whether the benefit is individual or collective.

- Socially acceptable use of data is not only a moral imperative, it is closely tied to reputational risk for partners in an open banking ecosystem.
- The use of financial data should be relatable and have daily relevance for the people whose data are shared.
- Helping people use their data for socially beneficial purposes can build bridges over the trust gap.
- Transparency and consent are not the only safeguards: public Data Protection Impact Assessments and data minimization can also build trust.
- People who do not trust the partners in an ecosystem, may trust the system itself if there are reliable standards and ways for them to seek recourse when something goes wrong.

Discussion and main takeaways

The roundtable yielded use cases for banking data in smart cities within three categories:

1. efficiency gains in administrative processes,
2. input for evidence-based policy making in economics and sustainability and
3. incentives for citizens towards positive sustainable behaviour.

For all participants in such an ecosystem, primarily for the citizen, there should be an added value of participating in the ecosystem. This includes financial institutions, (local) public authorities, FinTechs, other providers, and above all, the bank clients/citizens. This added value can be an individual or collective benefit, but it should in any case be clear and understandable to anyone involved.

Using open banking data in smart cities faces major challenges in ensuring adherence to data protection regulations and ethical frameworks, trust of the citizens, opening data for all the stakeholders and incomplete data. In order to overcome these challenges, the major solutions proposed include using technologies such as privacy enhancing technologies, and non-technological solutions such as educating citizens, creating transparency and introducing policies to stimulate the sharing of data in a privacy secured manner. Technology is evolving in the right direction with the Solid infrastructure, allowing citizens to be in control of their data and to decide themselves with whom it can be shared, based on the transparent added value that they will receive through this data sharing.

Citizens' trust and social license are closely tied to reputational risk for partners in an open banking ecosystem. Trust between partners in the ecosystem also needs to be given special attention. As for trust from the public, instead of relying on privacy notices, ecosystem partners need to engage in dialogue with citizens and communicate in a relatable way using ordinary language. To improve trust in the ecosystem for all involved in it, there should be independent oversight and ways for citizens and partners to seek recourse should anything go amiss.

Other takeaways are:

- Transparency means informing citizens about ecosystem partners, purposes, and risks.
- Consideration must be given to 'vulnerable groups', who may need more explanation.
- Respecting data minimization and publishing DPIAs also help improve trust.
- Helping people use their data for socially beneficial purposes can build trust.