On Data Privacy, Governance and Portability:

Turning Obstacles into Opportunities

Authors: Felix Wagner, Daniel Alarcon-Rubio,
Saulius Grigaliūnas, Daugailė Syrusaitė

Date: 04 January 2021

Contents

For any inquiries regarding the document, please contact:

Felix Wagner, Trafi employee, felix@trafi.com

Saulius Grigaliūnas, Senior Data Engineer, saulius@trafi.com

Daugaile Syrusaite, Lead Legal Officer, DPO, daugaile@trafi.com

Preface

We live in the digital age and each of us generates large amounts of data through the course of our daily lives simply by using the digital tools, private and public, at our disposal. In every single one of those tools, we leave digital footprints which are subject to be exploited, to a larger or smaller degree, for an individual, a collective public, a collective private or a third party's own advantage. More often than not, the individual is not in control, sometimes not even aware of the options and possibilities that such exploitations can bring forth. Because of that, regulations arise to return control, delimit exploitation and protect every individual's privacy and autonomy.

This kind of protection is of fundamental importance for a modern society to function, yet its current implementation misses to create a breeding ground to realise the positive possibilities that these data elements can bring to individuals and society. This is ever so prominent in the mobility space, when such space is enabled digitally (via for example, itinerary choices, purchases, trip executions, etc.) - we argue that there is a data governance void in the German smart mobility market which is incidentally fundamental for the success of solutions, such as Mobility as a Service. On top, the apparent lack of competitive technologies compliant with the prevailing regulations (such as the GDPR), leave few options for innovation for European-based businesses, as affordable suitable technologies tend to originate in the US.

Because of this, we would like to explore the opportunities that may arise as we put the end user, the data subject, at the centre of the control. We would like to investigate protection, portability and legal implications of such an approach as well as explore technological challenges, possible product features and financing models.

1. Introduction

Digitalization is shaping our everyday life and creating novel forms of living, travelling and working. The fuel driving this trend is data in its digital form, which is generated in an unprecedented manner. In fact, recent reports estimate that 2.5 quintillion bytes of new data are created every day, and this number only tends to increase as more people get access to the internet and more analog assets become digitized (Domo, 2020).

This trend can be quite overwhelming, especially for the individual. Data is constantly being generated, but its application, its location and its processors as well as their interests seem to be inscrutable. In European context, this lack of transparency is addressed through legislation that is very solid in theory, yet fails to keep up with the rapid pace of global technological development in practice. This leaves data generators in the dark, powerless and exposed to potential misuse and violation of their individual privacy. Such lack of transparency and control creates societal mistrust towards data-based solutions to such an extent that it may demolish the tremendous potential of data-based solutions in the future.

Nonetheless, disregarding digital technology cannot be the answer to complex socio-environmental systems such as urban transport networks. Due to climate change, rapid urban migration and environmental degradation, there is a drastic need for a transformation towards a sustainable, low-carbon mobility future - a future that enables participation, inclusion and economic prosperity alike (Creutzig, 2019). However, this transformation will only be possible if the tremendous potential of data-driven technology is utilized and the insights generated by analyzing large amounts of data are acted upon (Kitchin, 2015).

We deem individual privacy and data protection as a fundamental human right and we highly support the recent developments in the context of data protection, such as the released list of recommendations for additional measures for personal data transfer to countries outside of the EU by the European Data Protection Board (EDPB) (2020) or the proposal for a data governance act by the European Commission (2020b). Yet, we also recognize that further regulation might not be enough to fully address existing challenges. In fact, it might limit the innovation of data-based solutions even more. Instead, individual data protection, trust and agency will emerge through strong institutions and data governance, which builds on existing legislation and empowers the individual. Data literacy and transparency are needed to fully grasp the risks and opportunities associated with data. Usability and control are needed to translate complex problems into accurate and efficient decisions.

In the following sections, we will highlight the need for better data governance in the German smart mobility market, critically examine possible options to govern data, and propose a novel approach that limits the shortcomings of the existing governance model and focuses on empowering the individual.

2. Current Status of Affairs

2.1. Data - the Fuel that Drives Smart Mobility

Within the context of climate change, rapid migration and environmental degradation, urban systems are facing great challenges. Transport systems, for example, are responsible for providing access to work, education, healthcare, and sociocultural activities within a city (Umwelt Bundesamt, 2019). At the same time, these systems are charged with the task of drastically reducing emissions, improving spatial availability, raising passenger capacity and reinforcing security (Seto et al., 2014; Sze and Christiensen, 2017; European Commission, 2017).

In light of these challenges, the adoption of data-driven technology offers great potential for fostering sustainable urban mobility (Kitchin, 2015). Through data sensing and sharing, institutions gain access to richly detailed data, and analysis conducted via big data and machine-learning methods generate new insights benefitting transport operators, planners and users (Davidsson et al., 2016). From an economic perspective, those benefits include a higher level of efficiency within transport systems (Melo et al., 2017) as well as potentially lower costs for transport users (Cohen et al., 2016). In addition, smart mobility solutions can reduce the environmental impact of transport systems, as is the case with Mobility as a Service (MaaS), bike sharing or car sharing, which reduce reliance on privately owned cars. This can then lead to reduced congestion and transport-related emissions (Docherty et al., 2018), as well as more free space in cities (Creutzig et al., 2020). Data-driven mobility can also improve inclusion and participation within transport systems by providing tailored, on-demand services which satisfy individual mobility demand (Docherty et al., 2018). In addition, it can also incorporate user feedback for better service provision (TU Berlin, 2020).

Despite this tremendous potential, digitalization is frequently met with great uncertainty. This is because digitalization essentially redefines the concepts of power, security, privacy and autonomy (Zuboff, 2019). In order to reduce the associated risks and leverage the immense potential of data-driven technology, data governance is of utmost importance.

2.2. The Importance of Data Governance

According to the latest data strategy of the European Commission, data governance can be defined as “the organizational approaches and structures (both public and private) to enable data-driven innovation on the basis of the existing legal framework” (European Commission, 2020, p.8). At the core of data governance lie the challenges of enabling large-scale data sourcing, orchestrating data sharing between various stakeholders and creating a baseline of trust that allows digital innovation to prosper (Scassa, 2020).

Using the EU Commission’s latest data strategy (2020a) as a framework, data governance is considered to enable the following:

- Data availability
- Data interoperability and high quality levels
- Prevention of power imbalances
- Secure, efficient infrastructure
- Empowerment of individuals

As value can only be extracted from data if the data in question is accessible, data governance needs to ensure that data is made available (Abella et al., 2017). In the context of smart mobility, this includes opening up government-held data, especially as publicly funded processes produced this data (European Commission, 2020). This also includes the enabling of private sector companies to share data with other institutions without losing their competitive advantage (Docherty et al., 2018), as well as fully protecting the privacy and autonomy of individuals when sharing personal related data (Creutzig et al., 2019)

In addition to availability, data governance has to ensure interoperability between different data sets (achieved through coherent data standards) and high-quality data. Both enable the efficient sensing, storing, sharing and analysis of data. In contrast, non-user-friendly or proprietary data formats and inaccurate data require time spent cleaning data sets. In addition, they lead to a low quantity of usable data - or worse, to wrong decisions (Sadiq and Indulska, 2017).

On top of that, data governance has to counter digitalization-specific power concentrations arising from monopolistic market structures. Scholars like Docherty et al. (2018) argue that data (hence, power) monopolies can take advantage of their position and only pursue their own interests instead of serving the public good. Therefore, the involvement of the public and private sector has to be carefully defined. This also incorporates the potential use of alternative organizational and decision-making formats such as distributed ledger technologies (DTL) (Lopez and Farooq, 2018) or data trusts (Blankertz, 2020; Open Data Institute, 2019).

Data can only be processed within a functioning and secure infrastructure (such as cloud infrastructure). However, this poses another challenge for data governance in the context of smart mobility, because energy-efficient, cost-saving and secure data-processing capacities must also be present. This also includes reducing the existing reliance on outer European companies to provide data infrastructure (European Commission, 2020).

Lastly, data governance should empower individuals to have more control over their individual data. In the context of urban environments, scholars like Chyi and Panfil (2020) argue for more citizen participation in the process of defining the purposes of data use. This not only enhances transparency, knowledge and trust, but it also enables the development of better services, as initiatives like Posmoe Switzerland (2020) or SimRa at the TU Berlin (2019) are impressively demonstrating.

Case Study Example: Data Governance of Smart Mobility in Germany

When honing in on the smart mobility market in Germany, it becomes clear that the general data governance model follows a ‘regulated market’ approach, in which the sensing, sharing and processing of data is mainly regulated by a legal framework developed by the European Union and enacted by its individual member states. Legislation includes, for example, the free flow of non-personal data (FFD), the Open Data Directive, the General Data Protection Regulation (GDPR) as well as the Cybersecurity Act (European Commission, 2020). EU-based regulation is accompanied by additional legislation on the national level, such as the Act against Restraints of Competition (ARC) in Germany (Schoening and Ritz, 2019).

This regulatory framework ensures data availability in the German smart mobility market, as a number of cities such as Hamburg and Berlin are following the Open Data Directive and publishing infrastructure and non-personal data via so called open-data-platforms[1] (City of Hamburg, 2020; City of Berlin, 2020). In contrast, private sector companies or individuals are not obliged to share or publish their data.

[1] For more information about which cities are running an open data platform in Germany, please visit the Open Data Atlas at http://opendata.tursics.de/

Instead, they negotiate data sharing arrangements via contractual agreements based on the GDPR. Those sharing arrangements are then assessed via public data protection agencies such as the Berliner Commission for Data Protection and Freedom of Information (2020) (DPA) to ensure individual data protection.

Data interoperability has not been addressed in the 'regulated market' model thus far. This causes integration problems (Wagner, 2020), yet leaves room for the implementation of the newest data formats and standards in smart mobility applications. In contrast, power imbalances are currently being countered through regulation such as the ARC. In this case, it is important to note that while the German smart mobility market is still very fragmented, many transport companies provide digital services using APIs of outer EU-based tech companies like Google or Paypal. Similarly, a number of transport providers utilize data infrastructures to host their services, which are provided by outer EU-based companies such as Amazon or Microsoft. Finally, the empowerment of the individual is receiving little attention from both data processors and public institutions. So far only small research initiatives such as SimRa at the TU Berlin (2020) have tried to implement data sourcing, sharing and processing in line with clear transparency and user control policies.

2.3. A Case for Action on Advancing Data Sharing Frameworks

While the 'regulated market' governance model aims to leverage digital innovation, create equal conditions for competition and protect individuals' interests, there are a number of barriers in place currently preventing the smart mobility market from harnessing its full potential. On the one hand, these barriers are relevant to the topic of data protection and individual trust. On the other, they block data portability and innovation in the smart mobility domain.

2.3.1 Contractual agreements are not enough for data protection and value creation

In the context of smart mobility, transport users need to share their data to receive personalized, on-demand services. Hence, in order to enable data availability, data processors and data subjects negotiate sharing arrangements based on the rules defined in the GDPR. However, while the GDPR sets clear rules for use of personal data that function well on a theoretical level, it does not consider the fact that in all practicality, data-based solutions are built out of a complex interplay of multiple software applications from different institutions. Moreover, it fails to grasp that an even more complex flow of data arises out of this complex interplay. As a result, the data subject is left with either over-generalized or too specific, excessively time consuming consent choices when using smart mobility services. In other words: while technology rapidly scales to create unforeseen benefits out of data, data protection fails to keep up the pace, and the data subject has to suffer the consequences (Ullbricht and Pallas, 2020). These consequences include the data subject neither understanding where their individual data is located, processed and used, nor having any control over it. The only reasonable reaction to this state of vulnerability is either resignation (in the form of giving consent without reading data privacy statements) or to mistrust the institutions involved in providing smart mobility solutions and to minimize further data availability.

This mistrust is emphasized by the fact that many building blocks of smart mobility are provided by US based institutions and their usage requires personal data to be transferred to the US - a country with far less protective regulation on data privacy (Editorial Board, 2019).

In this context, however, we highly support the developments which have taken place since the Schrems II ruling. Especially, the released list of recommendations for supplementary measures when transferring personal data to institutions outside of the EU by the EDPB (2020) constitutes a great step towards better protecting every individual's privacy.

Nonetheless, it remains to be seen how companies will implement the proposed measures in practice, as the list of recommendations does not provide a "perfectly packaged solution[...] to the very real and practical challenges that companies face" (Fennessy, 2020). In addition, it has to be critically questioned whether the recommendations are enough to not only protect every individual's privacy but also to support data-based innovation. Considering the severe challenges urban transport systems currently face and taking into account the undisputed offering of US based services (as well as the sheer absence of competitive alternatives from the EU sector), it has to be interrogated whether it might need further data governance measures that sufficiently protect personal data, yet better integrate highly important services from companies located outside of the EU.

2.3.2 Data Portability will only be possible if trust and transparency are ensured

Similar to data protection, data portability of personal data is also determined in the GDPR. At its core, Article 20 of the GDPR aims to provide more control and autonomy to the data subject and to counter power imbalances in the digital realm. In practice, however, the article's unclear scope and lack of practicability can be criticized. This becomes clear when the data subject requests an archive of its user data as the shared data formats tend to lack the required semantics and standards to seamlessly migrate this data to another platform, leaving the data subject without any control and autonomy (Göndör, 2017). In addition, the data subject is left in the dark (again) as the right to portability does not necessarily apply to all cases of personal data (Drechsler, 2018). As a consequence, data is kept locked in small and large silos, giving only a few large tech companies the possibility to create value out of big data analysis. This value, however, has to be critically questioned as it is clearly guided by a few institutions' interests.

The EU has made great progress in addressing this challenge, by releasing its proposal for a data governance act (2020b) which aims to strengthen data intermediaries in order to facilitate better data interoperability between distributed data silos. Similar to the list of recommendations by the EDPB, however, its practical implementation and potential consequences remain uncertain (Dr2consultants, 2020). Yet, it has to be questioned whether the data governance act will be enough to sufficiently define the trade off between unlocking distributed data silos, fostering data-based innovation and sufficiently protecting an individual's privacy. Therefore, it is of utmost importance to examine what lies beyond the implementation of the data governance act.

To sum up, we argue that despite the recent positive actions undertaken by the EU, the current data governance model of the German smart mobility market has shortcomings. It not only fails to sufficiently protect the privacy of the individual and to create trust, it also risks not making use of the tremendous potential of digital solutions in the face of severe environmental, social and economic challenges.

To address these challenges, this whitepaper aims to take on a long-term view and focuses on how data governance can be designed to create a breeding ground for digital trust and innovation. As such, the following section proposes different data governance options and examines how sufficient they can address all five key areas of data governance. By doing so, it analyses to what extent the proposed solutions might be more adequate to define the difficult trade-off between creating value out of data, while ensuring solid data protection.

3. Possible Data Governance Solutions

To counter the examined challenges, we analyze different options that might be more sufficient in addressing the key areas of data governance. We adopt the option of 'doing nothing' as a baseline of comparison. The different governance options include:

- Doing Nothing
- Using Only Regulation
- Data Trust
- Distributed Ledger Technology (DLT)
- User Centred Trust (UCT)

Each of those options are evaluated according to their usefulness, when ensuring:

- Data Availability
- Data Interoperability and High Quality
- Prevention of Power Imbalances
- Secure and Efficient Infrastructure
- Empowerment of Individuals

Table 1 summarises the findings of the analysis:

Options / Tasks	**A. Data Availability**	**B. Data Interoperability**	**C. Power Imbalances**	**D. Data Infrastructure**	**E. Empowerment**
1. Doing Nothing	-	+/-	-	+/-	-
2. Using Only Regulation	+/-	+/-	+/-	+/-	+
3. Data Trust	+	+/-	+/-	+/-	+
4. DLT	+	+	+/-	-	+
5. UCT	+	+/-	+	+	+

Table 1: Analysis of different data governance options, where (+) indicates that the governance task is sufficiently addressed and (-) indicates that the governance task is not sufficiently addressed.

Data Governance Option 1: Doing Nothing

As seen in the previous chapter, the current data governance model has far-reaching consequences for individuals, businesses and governments aiming to utilize smart mobility solutions. Therefore, doing nothing would only reinforce a lack of data availability due to individual mistrust, which would then impede future innovation of smart solutions. With the adoption of the data governance act, interoperability will potentially be improved, yet its effectiveness in practice remains to be seen (Fennessy, 2020). Furthermore, power imbalances would continue to grow, as only a few large institutions would drive innovation based on the analysis of immense volumes of proprietary data, reinforcing their power. Data infrastructure would continue to be provided by outer EU-based institutions. This constitutes a sufficient solution from an environmental, security and cost perspective, yet it further strengthens these institutions' power and might reinforce high dependencies in the data infrastructure sector of the future. Finally, as a complete removal of smart solutions from the mobility sector is not a viable option, individuals would be left powerless and without transparency, living in a state best described by Draper and Turow (2019) as 'digital resignation'.

Data Governance Option 2: Using Only Regulation

Adopting more regulation might contribute to innovation because it can enhance the availability of government-held or proprietary data. This is the case with the Transport Act in Finland, for example, which sets clear rules for data sensing and sharing in the mobility sector (Ministry of Transport and Communication, 2018). Regulation also has the potential to improve data interoperability by creating clear standards for data sharing or, as addressed in the data governance act, by strengthening the trustworthiness of data intermediaries (European Commission, 2020b). Yet, the wrong type of regulation might also counter data availability, interoperability and innovation: over-protective standards, constant pressure for companies to be compliant as well as slowness and inflexibility applied in the context of highly agile technological development can all be hindrances to innovation. (Creutzig, 2020). In addition, it can take years until new regulation is passed, meaning it might already be outdated when it finally enters the market. Similarly, it can be useful to counter market imbalances by regulating monopolistic service providers in the European context. Yet this would be neglect of highly advanced services, which drastically conflicts with the goal of utilizing data-driven technology for the creation of public value. From a data-infrastructure perspective, this governance solution does not necessitate new, unsecure or energy-intensive infrastructure, yet regulation of existing data infrastructure providers might also imply neglecting highly advanced services. Nonetheless, applying regulation could be a way of sufficiently protecting individuals' privacy and autonomy.

Data Governance Option 3: Data Trust

The third option implies the usage of a data trust framework. In such a framework, an independent third-party institution without vested financial interest in the value derived from data ensures that data is only used according to predefined purposes collectively agreed upon by all stakeholders (Open Data Institute, 2019).

This raises the trust upon which stakeholders (including individuals) might be more willing to share data and contribute to data availability.

While its impact on data interoperability remains unclear, cooperation via the trust might facilitate agreements on collectively used standards and interfaces. Data trusts also have the potential to effectively counter power imbalances as the governance of data is decoupled from monetary or power-related interests (Creutzig, 2020). However, it is important to note that in the context of a lack of competitive alternatives from the EU, even a data trust is not able to prevent the use of services from monopolistic third-party providers located outside of EU jurisdiction. Furthermore, while it does not require additional energy-intensive or unsecure data infrastructure, the underlying business model, complex consent mechanisms and financing are yet to be defined, making it currently inapplicable for practitioners (Open Data Institute, 2019). Finally, as seen from pilots conducted by the Open Data Institute (2019), data trusts can work in favor of data subjects' needs and sufficiently include them in defining the use of personal data.

Data Governance Option 4: DLT

DLT, such as blockchain technology implemented in a MaaS application as proposed, for example, by Lopez and Farooq (2018), could heavily contribute to data availability and interoperability, as all participants are equally powerful and all data transactions are transparent and secure. DLT received huge attention in the context of MaaS, as it has the potential to reduce power imbalances arising from an intermediary data-collecting institution by leveraging decentralized consent mechanisms. In Germany, however, MaaS data controllers are usually public transport providers, meaning institutions controlled by public authorities. We therefore see little potential for an abuse of power arising from this central position. Instead, power imbalances in the smart mobility market arise through third-party services that are currently irreplaceable, but they are also provided by already very powerful institutions that are constantly gaining more data and leveraging network effects. Furthermore, DLT addresses mobility needs with market-based mechanisms, where supply meets demand and vice versa; but it is highly questionable as to whether such liberal market mechanisms are sufficient to fully address every aspect of a highly complex, socio-environmental system such as a transport system. In addition, while DLT proves to be a very secure approach, it is highly criticized for its inefficiencies, lack of scalability and high-energy usage (Truby, 2018). Nonetheless, data subjects could receive a high degree of power, as they would directly sign so-called 'smart contracts' with the institution providing a service to them. Data transfers would therefore be very transparent and data subjects would be in complete control over their data.

Data Governance Option 5: UCT

UCTs follow a human-centric approach in which the individual has full control over their personal data. They utilize a privacy-by-design default paradigm with state-of-the-art encryption and pseudonymization methods in which decryption keys are stored in an independent third-party institution. In addition, the responsibility to decide who receives access to decryption keys is handed over to the data subject. By providing transparency and control, a foundation of trust emerges which has the potential to strongly contribute to data availability. While it does not necessarily contribute to coherent data formats - in other words, interoperability in the smart mobility market - it enables the reuse and sharing of data between institutions and for different purposes according to individual consent.

Furthermore, it is very effective in countering market imbalances that arise from the concentration of large amounts of data at a few institutions: even when using such services, the control over the data stays in the hands of the data subject. Similarly, it enables the use of efficient, secure and cost-effective data infrastructure provided by companies outside of the EU while ensuring sufficient protection of a data subject's privacy.

Finally, it strongly contributes to the empowerment of individuals, as it enables transparency and provides full control over personal data.

In summary, our analysis demonstrates that more-efficient data governance options than ‘doing nothing’ do exist. It can be seen that ‘Using Only Regulation’ might be sufficient to fill some areas of the current data governance void, yet it appears to not be enough to enable both data protection and innovation. In contrast, emerging concepts such as ‘Data Trusts’ or ‘DLT’ seem to be more sufficient in solving complex governance questions but lack practical feasibility. As a consequence, we recognise UCTs as having the highest potential to sufficiently address the current data governance void. The following section will thus analyze UCTs in more detail.

4. Deep Dive: User-Centred-Trust Framework

The fast developments since the Schrems II ruling make it extraordinarily difficult for companies operating in the digital realm to be compliant and avoid potential fines. Yet, we argue that striving to be compliant is not enough to drive change towards transparent, fair and prosperous data ecosystems - ecosystems that constitute breeding grounds for technological innovation and value creation.

When faced with unparalleled challenges within urban environments, a data-governance approach that sufficiently protects each and every individual yet also fosters the creation of innovative data-based solutions is required. Building upon the results of Chapter 3, we propose UCTs as an effective data governance solution. This novel approach reflects the high importance of trusted intermediaries, as highlighted in the proposal for a data governance act (EU Commission, 2020b), and grounds its core principles on the recommendations for additional technical measures by the EDPB (2020). In addition, it recognizes 1) the high importance of trust in the context of uncertainty, and 2) the tremendous potential that lies within the use of state-of-the-art data protection technology. The following examines each component of UCTs - trust and technology - in more detail.

4.1 From digital resignation towards empowerment and innovation - the importance of trust in the digital age

Social science research has shown that data subjects currently live in a state of 'digital resignation', where they "desire to control the information digital entities have about them but feel unable to do so" (Draper and Turow, 2019, p.1). Stemming from this helplessness, the rational response to digital technology is mistrust, which according to Steedman et al. (2020) is a complex phenomenon occurring both generally and on a local, personal level. In this context, 'generally' refers to the data ecosystem as a whole and the ambiguous flow of data between different institutions, while the 'local' concerns the individual's experience and control in the context of individual data protection.

The concept of UCT addresses both of these dimensions. General mistrust towards data-based solutions is countered by utilizing state-of-the-art encryption and pseudonymization techniques to ensure data protection across all layers of a data ecosystem (see chapter 4.3.4). In addition to this, UCT is aimed at ensuring full transparency and comprehensibility on the flow of personal data (as in, between which institutions data is being shared, or for what purpose is it shared; see chapter 4.3.2). Mistrust on a local level can then be countered by empowering the individual and handing over full control to the data subject (see chapter 4.3.3). Building upon the findings of prior projects, this individual autonomy also aligns with the fact that "[t]he benefits and risks of data sharing may well be seen very differently by different groups" (Burall et al., 2019, p.6).

As a consequence, a new foundation of trust emerges that can be utilized to generate high value out of data for all stakeholders involved. Prior projects that incorporated a similar aim to move beyond 'data protection for compliance' have impressively demonstrated the high potential of trusted ecosystems for data-based innovation. Two such projects are described in Box 1 and Box 2 in more detail.

Box 1: Using synthetic data sets to foster innovation in the healthcare sector.

HEALTHCARE

The benefits of data for the healthcare industry are undisputed. Big data can foster highly personalised treatments, enhanced clinical efficiency, more efficient medical supply chains and advanced medical research (Singhal and Carlton, 2019). Yet they also necessitate high data protection standards, as personal healthcare data is highly sensitive. Innovations such as the Synthetic Data Engine from the private company Mdclone go beyond compliance and enable the sharing and analysis of personal data through synthesising and randomizing data sets, detaching personal information while retaining statistical properties (Mdclone, 2020).

Box 2: Utilising crowd sourced data to improve safety in the transport sector.

TRANSPORT

In the research project SimRa at the TU Berlin, bike riders in Berlin voluntarily collected data about daily routes and potentially risky locations such as inadequate bike lanes or unclear routing. On top of existing legislation, this data is locally pseudonymized ('on-device') and shared only according to individual consent. Examples of partners for data sharing include local authorities who can use this data to improve the city's bicycle infrastructure and encourage cycling, thereby reducing air and noise pollution and traffic congestion (TU Berlin, 2019).

4.2 Technology that upholds trust is the key to unlock the potential of data

Building upon the findings in the fields of social science, technology becomes a vital factor in enabling trusted data ecosystems. The following section analyzes the changing structure of data ecosystems when implementing UCTs, defines the consent mechanism, and illustrates its implementation via encryption and pseudonymization methods.

4.2.1 Schematic description of a UCT data ecosystem

In a traditional setup without a UCT, the data controller is located in the center of the data ecosystem of an application and receives data from a data subject. This data subject also shares revenue with it to receive a service. The data is then transferred via a data processor[2] to a third-party data user who provides a service (i.e., storing or analysing the data). The value created by the combination of different services is then returned to the data subject.

In contrast, a UCT places the data subject in the centre of the data ecosystem. As soon as the data subject shares data with the data controller, it becomes encrypted and the keys are stored in a trusted, independent third-party institution, the so-called 'ledger'. Once again, the encrypted data is transferred via a data processor to third-party data users who provide a service in order to create value from data.

[2] Note: the data processor is not displayed in Figure 1

This time, however, the data subject has the power to control which third-party institution processes its data. By sharing a 'consent policy' with the ledger, the data subject chooses whom the ledger shares the appropriate encryption keys with in order to decrypt and process the data. Figure 1 schematically compares the traditional data ecosystem (left) with an UCT (right).

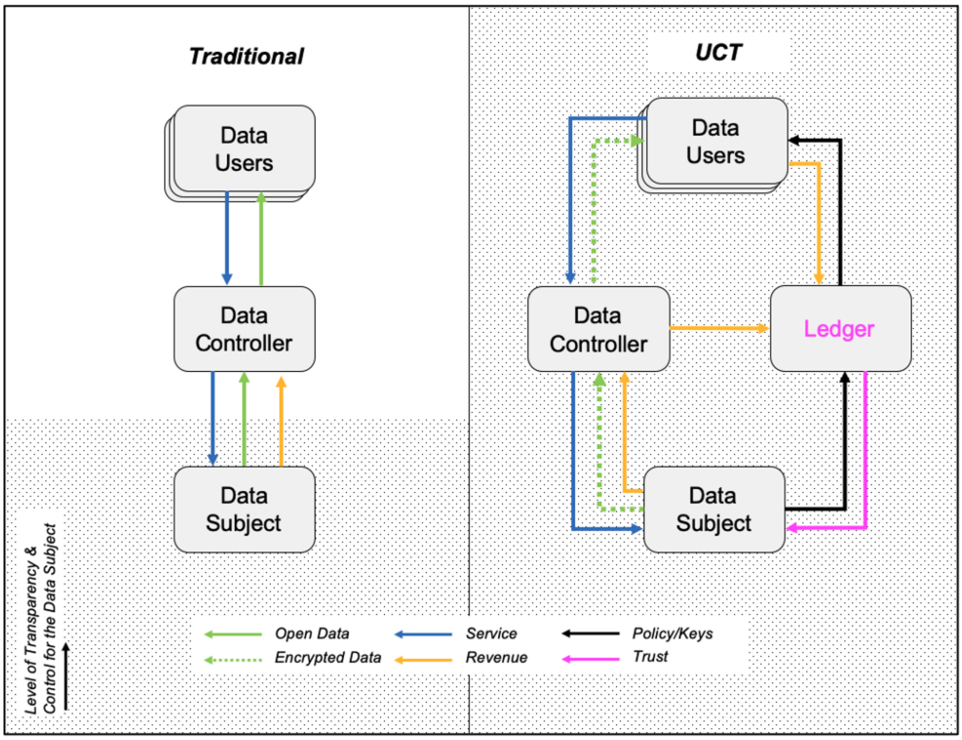


Figure 1: Traditional data ecosystems compared to UCT data ecosystems. The arrows mark the flow of data (green), services (blue), revenue (orange), consent policy/encryption keys (black) and trust (pink). The dotted background colouring marks the area of transparency and control for the data subject.

4.2.2 Changing Roles and Responsibilities in a UCT

With the adoption of UCTs, existing roles and responsibilities change and new ones are introduced. Starting at the bottom of Figure 1, the data subject receives a new role, as they have more transparency and control, and by extension, have more responsibility to govern their data. The ledger, which is an independent third-party institution with no financial interest in the direct value created out of data, is introduced next to the data subject and aims to serve as trusted intermediary between the data subject and the different data users. By doing so, it reflects one of the core ideas of the latest proposal for a data governance act by the EU commission (2020b). The ledger's responsibilities will include efficiently translating consent policies into effective key sharing as well as operationalizing effective key management. The roles of the data controller and the data users will stay unaltered, but they will become more dependent on the individual decisions of the data subject. Finally, as all actors in the UCT will strongly benefit from more available data, transparency and trust, they will become responsible for financing the ledger.

Potential financing models include (sorted from high to low probability):

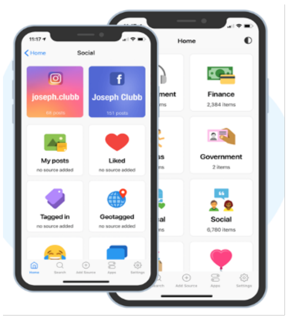
- Taxing data transfers for data users and data controllers (as displayed in Figure 1)
- Subsidizing the trust as part of the public transport system
- Financing the trust as part of the public infrastructure
- Financing the trust via the data subject

4.2.3 Operationalizing a Consent Policy

As the GDPR constitutes the core foundation for every action on personal data, UCTs are also grounded on the basis of the requirements set out in the GDPR. We recognise that lawful and trustworthy data transfers can only be ensured if “the data subject has given consent to the processing of his or her personal data for one or more specific purposes” (Article 5, Regulation (EU) 2016/679). Furthermore, the GDPR recognises that consent, in order to be valid, must be:

- **“freely given”** (Article 4, Regulation (EU) 2016/679)
- “a **specific**, **informed** and **unambiguous** indication of the data subject’s wishes” provided “by a statement or by a clear affirmative action” (Article 4, Regulation (EU) 2016/679)
- “easily **retractable**” (Article 7, Regulation (EU) 2016/679)

In order to operationalize an effective consent mechanism, the data subject must then be (a) thoroughly informed and (b) able to make specific, unambiguous and retractable choices. While this effectively translates the legal framework into action, we recognize achieving comprehensible and accurate consent requires an understandable and user-friendly design. We therefore propose the following:



(a) A dashboard designed to inform the data subject, illustrating which data has been shared with which institution and for what purpose. At the same time, this dashboard should be used to highlight the value of additional data sharing. For example:

Purpose: Analytical purpose for research in the field of sustainable transportation.
Institution: TU Berlin
Data: Location Data, Mode Choice, Gender, Age

Figure 2: Exemplary Dashboard from the App ‘digi.me’ (2020)

(b) To be able to make free, specific, unambiguous, retractable choices, we propose a layered, purpose-based consent workflow. Here, the data subject can select for which general purposes they agree to share personal data. However, if the data subject wants to be more specific, they can also allow data sharing on a more granular level and select purposes or institutions. As displayed in Figure 3, the data subject can, for example, allow ‘Additional’ data sharing for ‘Research’ and ‘Innovation’ purposes, but they can also prevent sharing their personal data with local institutions from the “Community”, such as charities or non-governmental organizations.

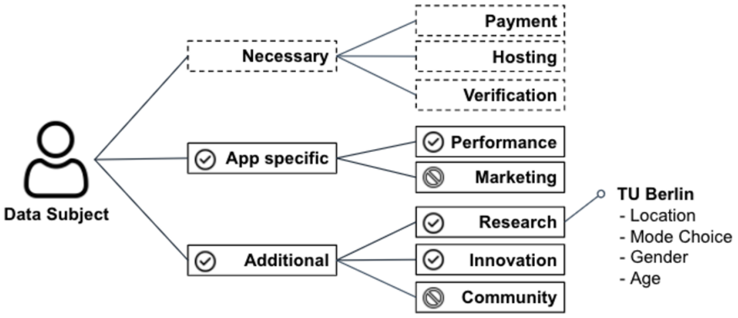


Figure 3: Layered, purpose-based consent workflow (Continuous border linestyle marks consent options, dashed linestyle marks necessary functions which cannot be opted out).

4.2.4. Protecting Data with Technology

Finally, to enable this kind of control for the data subject, we propose utilizing state-of-the-art data protection technology as well as privacy-by-design and default principles. These measures reflect the latest recommendations for additional technical measures by the EDPB (2020). The following summarises the approach:

a. Protection of 'Static Data'

- AES-256 encryption algorithm (leading standard in bank and healthcare sector; optional AES-128 for low-power devices).
- Files will be encrypted individually to prevent one key being cracked and the opening of all data files.
- Files will be encrypted individually in addition to encryption by storage providers.

b. Protection of 'Data in Flow':

- Data transmissions will only be possible via encrypted connections with authentication of receiver and sender.
- Avoidance of 'man in the middle' through transit validation.
- Data will be encrypted as soon as it enters the system.

c. Protection of 'Data being Processed':

- When data is transferred to third countries for processing, pseudonymization will be utilized before the data is transferred. By doing so, personal information cannot be attributed back to the data subject, yet remaining valuable information can still be processed.

d. Authentication & Authorisation:

- All system components that process personal data must be able to prove their origin (via a cryptographic certificate, issued by the ledger).
- All involved parties must be uniquely identifiable.
- User access must be authenticated (users must prove that they have the master password to define the consent policy).

e. Key Management

- There will be different keys for different purposes (for example, to avoid legal conflicts, local authorities must be able to obtain user data).
- The key lifecycle (from generation to deletion) will be determined in accordance with all involved stakeholders.
- There are already several third-party key management providers and adoption by major public cloud providers is beginning.

5. Conclusion and Next Steps

To conclude, this white paper recognizes that there is a data governance void within the German smart mobility market which diminishes the high potential of data and leaves individuals powerless and uninformed. To counter this, different data governance options were analyzed and the **User-Centred-Trust**, a novel framework that builds up on the latest recommendations of the EDPB (2020) and the proposal for a data governance act by the EU Commission (2020b), was chosen as the most promising solution.

To summarise, UCTs leverage:

- a **dashboard** that transparently visualizes data flow, all institutions involved and the purposes of data sharing
- a **layered, purpose-based consent workflow** to provide finely tuned control to the data subject
- state of the art **encryption** and **pseudonymization** methods to specifically restrict or allow access according to given consent policies.

With UCTs, a new paradigm can emerge in which individual data protection will go hand-in-hand with digital innovation. As UCTs have the potential to reduce existing friction in data ecosystems and increase the velocity of data sharing, unique innovations will emerge and value can be created not only for individuals but for society as a whole. This will unlock new ways of utilizing personal data, as data subjects can freely choose to share their data with other services, donate their data to support research, improve their digital experience or contribute to their community.

To implement UCTs in practice, the following key areas need to be addressed in the future:

- An independent third-party institution has to be determined that can act as a trusted ledger. Their roles and responsibilities, financing structure, legal implications, audit options and potential technology partners need to be defined in cooperation with local transport and data protection authorities. In addition, it has to be examined to what extent the trusted ledger can function as a trusted intermediary as defined in the proposal for a data governance act by the EU Commission (2020b).

- The dashboard as well as the layered, purpose-based consent workflow have to be designed, implemented and sufficiently tested. This will also include research on how the data subject's awareness of the importance of data protection and innovation can be gained.

- The encryption workflow and its requirements need to be developed in line with existing systems. This will include choosing a coherent framework for protecting static data, data in flow, and data being processed, as well as sufficiently testing the system, designing a coherent key management approach and choosing potential external partners.

6. References

Abella, A. et al. (2017) A model for the analysis of data-driven innovation and value generation in smart cities' ecosystems. Cities. 6447–53.

Berliner Beauftragte für Datenschutz und Informationsfreiheit (2020) Datenschutz und Informationsfreiheit. Jahresbericht 2019. [online]. Available from: https://www.datenschutz-berlin.de/fileadmin/user_upload/pdf/publikationen/jahresbericht/BlnBDI-Jahresbericht-2019-Web.pdf (Accessed 29 August 2020).

BfDI (2020) The BfDI's statement on the Schrems II judgment of the ECJ. [online]. Available from: https://www.bfdi.bund.de/EN/Home/Press_Release/2020/17_Schrems-II-ECJ.html (Accessed 10 November 2020).

Blankertz, A. (2020) Designing Data Trusts. Why we need to test computer data trusts now.

Burall, S. et al. (2019) GLA Greenwich Data Trust Pilot. Decision-Making Process. [online]. Available from: https://docs.google.com/document/d/1OiWdxSistGMqEEoez07BGUxpOEUiEWiqlVmtc9-iPQ8/edit#heading=h.gjdgxs (Accessed 10 November 2020).

Chyi, N. & Pan, Y. (2020) A Commons Approach to Smart City Data Governance. New America. 24.

City of Berlin (2020) Offene Daten lesbar für Mensch und Maschine. Das ist das Ziel. [online]. Available from: https://daten.berlin.de (Accessed 27 October 2020).

City of Hamburg (2020) Transparenzportal Hamburg. [online]. Available from: http://transparenz.hamburg.de/open-data/ (Accessed 27 October 2020).

Cohen, P. et al. (2016) Using Big Data to Estimate Consumer Surplus: The Case of Uber. National Bureau of Economic Research.

Creutzig, F. (2020) An integrated data platform to leverage the benefits of smart mobility.

Creutzig, F. et al. (2020) Fair street space allocation: ethical principles and empirical insights. Transport Reviews. 40 (6), 711–733.

Creutzig, F. et al. (2019) Leveraging digitalization for sustainability in urban transport. Global Sustainability.

Davidsson, P. et al. (2016) The fourth wave of digitalization and public transport: Opportunities and challenges. Sustainability. 8 (12), 1248.

digi.me (2020) digi.me Homepage. [online]. Available from: https://digi.me (Accessed 16 November 2020).

Docherty, I. et al. (2018) The governance of smart mobility. Transportation Research Part A: Policy and Practice.

Domo (2020) Data Never Sleeps 6.0. [online]. Available from: https://www.domo.com/solution/data-never-sleeps-6 (Accessed 27 October 2020).

Dr2consultants (2020) Data Governance Act: main elements and business implications [online] Available from: https://dr2consultants.eu/data-governance-act-main-elements-and-business-implications/ (Accessed 10.12.2020)

Draper, N. A. & Turow, J. (2019) The corporate cultivation of digital resignation. New Media & Society. 21 (8), 1824–1839.

Drechsler, L. (2018) Pratical Challenges to Data Portability in the Collaborative Economy. 20.

Editorial Board (2019) Why Is America So Far Behind Europe on Digital Privacy? New York Times. [online]. Available from: https://www.nytimes.com/2019/06/08/opinion/sunday/privacy-congress-facebook-google.html (Accessed 29 August 2020).

European Data Protection Board (2020) Recommendations 01/2020 on measures that supplement transfer tools to ensure compliance with the EU level of protection of personal data. Adopted on 10 November 2020. [online] Available from: https://edpb.europa.eu/sites/edpb/files/consultation/edpb_recommendations_202001_supplementarymeasurestransferstools_en.pdf (Accessed 10.12.2020)

European Commission (2020a) Communication from the Commission to the European Parliament, the Council, the European Economic and Social Comittee and The Committee of the regions. A European data strategy.

European Commission (2020b) Proposal for a Regulation of the European Parliament and on the council on European data governance (Data Governance Act)

European Commission (2017) European Urban Mobility Policy Context.

Facebook (2020) Updating our international data transfer mechanisms. [online]. Available from: https://www.facebook.com/business/news/updating-our-international-data-transfer-mechanisms/ (Accessed 10 November 2020).

Fazlioglu, M. (2020) Using SCCs post-'Schrems II': Guidance from DPAs. [online]. Available from: https://iapp.org/news/a/using-sccs-post-schrems-ii-guidance-from-dpas/ (Accessed 10 November 2020).

Fennessy, C. (2020) A breakdown of EDPB's recommendations for data transfers post 'Schrems II' [online] Available from: https://iapp.org/news/a/a-break-down-of-edpbs-recommendations-for-data-transfers-post-schrems-ii/ (Accessed 10.12.2020)

Göndör, S. (2017) The Importance of Data Portability and Interoperability in the Social Web. Telekom Innovation Laboratories TU Berlin.

Google (2020) Update to Standard Contractual Clauses (SCCs) (August 2020). [online]. Available from: https://support.google.com/adspolicy/answer/10042247?hl=en (Accessed 10 November 2020).

Kitchin, R. (2015) Making sense of smart cities: addressing present shortcomings. Cambridge Journal of Regions, Economy and Society. 8 (1), 131–136.

Lopez, D. & Farooq, B. (2018) A blockchain framework for smart mobility. IEEE International Smart Cities Conference 2018.

Mdclone (2020) Unlock Healthcare Data. Transform Care. [online]. Available from: https://www.mdclone.com (Accessed 10 November 2020).

Melo, S. et al. (2017) Guiding cities to pursue a smart mobility paradigm: An example from vehicle routing guidance and its traffic and operational effects. Research in Transportation Economics. 6524–33.

Ministry of Transport and Communication (2018) Act on Transport Services. [online]. Available from: https://www.finlex.fi/fi/laki/kaannokset/2017/en20170320_20180731.pdf (Accessed 28 September 2020).

MyData (2020) MyData [online]. Available from: https://mydata.org (Accessed 7 September 2020).

Open Data Institute (2019) Data trusts: lessons from three pilots.

Posmo Schweiz (2020) Genossenschaft POSMO Schweiz [online]. Available from: https://posmo.coop (Accessed 21 July 2020).

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (GDPR) (2016) Official Journal of the European Union L119, p.1-88/1

Sadiq, S. & Indulska, M. (2017) Open data: Quality over quantity. International Journal of Information Management. 37 (3), 150–154.

Scassa, T. (2020) Designing Data Governance for Data Sharing: Lessons from Sidewalk Toronto. Governing Data as a Resource.

Schoening, F. & Ritz, C. (2019) Germany's proposed digital antitrust law: an ambitious project to regulate digital markets. [online]. Available from: https://www.hlregulation.com/2019/10/31/germanys-proposed-digital-antitrust-law-an-ambitious-project-to-regulate-digital-markets/ (Accessed 10 November 2020).

Seto, K. C. et al. (2014) 'Human Settlements, Infrastructure, and Spatial Planning', in Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the

Intergovernmental Panel on Climate Change. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. pp. 923–1000.

Singhal, S. & Carlton, S. (2019) The era of exponential improvement in healthcare? [online]. Available from: https://www.mckinsey.com/~/media/McKinsey/Industries/Healthcare%20Systems%20and%20Services/Our%20Insights/The%20era%20of%20exponential%20improvement%20in%20healthcare/The-era-of-expontential-improvement-in-healthcare.pdf (Accessed 10 November 2020).

Steedman, R. et al. (2020) Complex ecologies of trust in data practices and data-driven systems. Information, Communication & Society. 23 (6), 817–832.

Sze, N. N. & Christensen, K. M. (2017) Access to urban transportation system for individuals with disabilities. IATSS Research. 41 (2), 66–73.

Truby, J. (2018) Decarbonizing Bitcoin: Law and policy choices for reducing the energy consumption of Blockchain technologies and digital currencies. Energy Research & Social Science. 44399–410.

TU Berlin (2020) SimRa: Sicherheit im Radverkehr [online]. Available from: https://www.digital-future.berlin/forschung/projekte/simra/ (Accessed 23 July 2020).

Ulbricht, M.-R. & Pallas, F. (2018) 'YaPPL - A Lightweight Privacy Preference Language for Legally Sufficient and Automated Consent Provision in IoT Scenarios', in Joaquin Garcia-Alfaro et al. (eds.) Data Privacy Management, Cryptocurrencies and Blockchain Technology. Cham: Springer International Publishing. pp. 329–344.

Umwelt Bundesamt (2020) Nachhaltige Mobilität. [online]. Available from: https://www.umweltbundesamt.de/themen/verkehr-laerm/nachhaltige-mobilitaet (Accessed 29 February 2020).

Wagner, F. (2020) Using Data to Foster Sustainable Mobility.

Zuboff, S. (2019) The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. New York: PublicAffairs.