MAX VON GRAFENSTEIN

How to build data-driven innovation projects at large with data protection by design

A scientific-legal Data Protection Impact Assessment with respect to a hypothetical Smart City scenario in Berlin

HIIG DISCUSSION PAPER SERIES 2020-03

ABSTRACT

As part of the research project "Data Protection by Design in Smart Cities"[1], this Discussion Paper consists of a legal-scientific Data Protection Impact Assessment (DPIA) discusses, on the basis of a technological Smart City project in Berlin (at Ernst-Reuter-Platz), how to build large data-driven innovation projects using a data protection by design strategy. The aim of this evaluation is to demonstrate how a DPIA can help to define a data protection by design strategy ensuring that the project meets the legal and societal expectations. The study illustrates that the EU General Data Protection Regulation does not forbid data-driven innovation projects *per se*, but rather forces the stakeholders involved to coordinate in due time *how* to process personal data to avoid unnecessary risks to individuals and the society as a whole. By means of certification schemes and codes of conduct, private companies can use such a data protection by design strategy as a competitive advantage and at least as a business opportunity. However, the study also points out some to-do's for the regulator: First, this study provides some conceptual clarifications on the methodology of the risk assessment, which are addressed to the Commission nationale de l'informatique et des libertés (CNIL) and the Technology Working Group ("AK Technik") of the Conference of the German Data Protection Authorities. Second, the regulator should follow its instruction from Art. 40 and 42 GDPR to support micro, small and medium-sized companies to set up certification schemes and codes of conduct for their processing activities. Third, national legislators should use the flexibility clauses in the GDPR to increase legal certainty for data-driven innovation projects, especially in the public interest, by applying a data protection by design strategy as presented. Finally, the EU legislator should reconsider its envisaged approach for the ePrivacy Regulation: instead of focusing on the consent of data subjects as the main legal basis for data processing, the legislator should establish a "legitimate interests"-clause *with an obligation to adhere to* an appropriate code of conduct or certificate for the processing activity in question.

KEYWORDS

Data-driven innovation, Smart City, GDPR, data protection impact assessment, data protection by design, ePrivacy Regulation, research purposes, Wi-Fi data, CCTV data, sensor data

AUTHOR INFO / AFFILIATION / FUNDING ETC.

Dr. Max von Grafenstein, LL.M., co-head of the research program "Governance of Data-Driven Innovation" at Alexander von Humboldt Institute for Internet and Society, and professor for "Digital Self-Determination" at Einstein Center Digital Future appointed to the Berlin University of the Arts.

For the conduct of this study, I would especially like to thank my colleagues Jörg Pohle, Jan Schallaböck, Kevin Klug, Jan-Philipp Siebold, and Mareike Lisker for their helpful comments and contributions.

This Discussion Paper is a result of the research project "Data Protection by Design in Smart Cities", which has been sponsored by Cisco Systems GmbH.

[1] See https://www.hiig.de/en/project/privacy-by-design-in-smart-cities/, last access: 6th May 2020.

PREFACE

This scientific-legal DPIA was mainly conducted in the second half of 2017. The time gap between then and its publication in the second quarter of 2020 is due, on the one hand, to the circumstances surrounding the application of the General Data Protection Regulation (GDPR) in May 2018 and, on the other hand, to the complexity of the subject-matter. When the first draft was finished, the team asked the financial sponsor of the research project, i.e. Cisco, to review several technical aspects presented in this DPIA. However, six months before the GDPR came into force, the contact persons on the sponsor side were so busy preparing for the GDPR that they were unable to read the draft and check the technical facts until shortly before 25 May 2018. Ironically, after the sponsor was able to review and approve the technical facts, the research team forwarded the draft to the Berlin Data Protection Authority (DPA). However, after the GDPR had become applicable, the DPAs were also so busy enforcing the GDPR that it was very difficult to obtain their feedback. (In this regard, it is important to note that we asked the DPA only for their expert feedback, not for an official statement; therefore the final findings and recommendations of this DPIA should be understood as our own position and not that of the DPA.) The final workshop with all representatives of the sponsor side and the DPA finally took place on November 2018. In view of the findings of the workshop, however, it became clear that the research team had to focus on the question of how the technical layer, the process layer, and the rule-making layer had to be designed precisely to become GDPR-compliant. Therefore, to explore in more detail how such layers can look like in practice, the research team initiated a second research project called Data Governance. In the course of this second project, it became clear that it was by no means absolutely necessary to set up the system in such a way that all data was stored immediately and centrally well in advance for some potential research purposes. This presumption (i.e. that the data is stored in advance and centrally) was the instinctive reaction of most experts who were asked for feedback in connection to the preceding Data Protection by Design in Smart Cities project and had therefore considered the Smart City project to be by no means GDPR-compatible. In contrast, the research team felt that there was at least a theoretical way to make such a system GDPR-compliant. This could, in short, be the case if the system was designed in a decentralized way and a trusted third party controlled not only the collection of data but also the use of the data. Even if some readers may not share this particular consideration – or further considerations – the publication of this DPIA will of course fulfil its main purpose: to start and promote an informed discussion among both experts and the general public on whether and, if so, how the Smart City system presented here could and should be designed, set up and operated.

So, the team started to revise its original draft to finally publish it in 2020. Surprisingly, almost nothing has changed since 2017: the ePrivacy Regulation is still under negotiation; the ECJ has even adopted the criteria proposed by the EDPB to determine when information relates to an individual; only on the methodical level, in particular in the conduct of a data protection risk assessment, has there been significant progress. However, in particular in this last respect, one may fear whether the current evaluation of the GDPR leads to changes within the law so that all methodical questions start all over again. To incorporate some of the findings of this DPIA into this GDPR evaluation process, and into the negotiation process of the ePrivacy Regulation, this DPIA had hence to be published. Ironically, the finalisation process was stopped again: the reason for this was that all Berlin libraries were closed due to the COVID-19 pandemics, just as the research team was about to finally check all footnotes for their topicality. However, the team decided not to wait any longer and could only check footnotes to the extent that they had online access to the respective repositories. Where this was not the case, this was noted in the footnote. Last but not least, some websites, which we accessed in 2017, no longer show all content. If we nevertheless present such outdated content or link to it (mainly to illustrate our theoretical reasoning), we also make this clear.

EXECUTIVE SUMMARY / RECOMMENDATION

SUMMARY

The present Data Protection Impact Assessment (DPIA) is a **combination of a scientific and a legal Data Protection Impact Assessment ("DPIA")**. Because the target of evaluation, i.e. the technical system in which the data is intended to be collected, is not yet sufficiently specified, this assessment cannot meet the requirements of Art. 35 GDPR. However, the current DPIA can serve as a productive basis for defining the data protection and security by design strategy for the development of the system. Although not required by law, such a scientific and legal DPIA offers two significant advantages:

- First, the DPIA helps to design the system at a **sufficiently early stage**, whereas if a DPIA had been conducted at the time when the system was already defined, changes to the system transposing certain data protection by design requirements would require more efforts than at an early stage of development.

- Second, by focusing on the *collection* of the data, the DPIA reduces the complexity of DPIAs that are necessary for applications that build on the *use* of the data. This helps **micro, small and medium-sized companies** focusing on data use because they can build on this DPIA and concentrate on the risks caused by their intended use.

After a description of the subject-matter in the first part of this DPIA, the second part examines the conditions under which the system could comply **with data protection laws**. This second part focuses on the collection of the data for research and statistical purposes in an urban traffic management environment. The following three findings are particularly relevant:

- All collected data is considered as **personal data** (i.e. Wi-Fi data, CCTV camera data, and also the parking lot sensor data), even if some involved actors seem to consider the data as "non-personal" or "anonymised". The reason for this broad approach is that as long as the data (just indirectly) relates to *an* individual (irrespective of who this exactly is), it could be used against this individual, for example, in the course of a legal procedure (e.g. by the police or insurance companies based on extra knowledge of a witness on-site). However, one can take the differences of *how* data relates to individuals into account to define the right data protection by design strategy.

- In view of the ongoing legislation process for the **ePrivacy Regulation**, the DPIA has not yet been able to clarify whether certain data types (particularly the Wi-Fi data) are covered by this regulation. If this were to be the case and the individual's consent would be required, this would significantly affect the research purposes (see in the next-following point why).

- As long as the GDPR applies, the collected data had to be based on the "legitimate interests"-clause to make sure that *all* collected data could be used for research and statistical purposes. The reason for this is that the data is collected *before* the data subjects can give their consent. However, this DPIA concludes that the data processing could be carried out (at least) for research and statistical purposes in an urban traffic management environment on the basis of the "legitimate interests"-clause, provided that (in particular) the following guarantees are met:

    - **Anonymisation or, more likely, pseudonymisation** of the data, in particular the Wi-Fi data through hashing and renewing the hash-values regularly, according to time intervals and locations that have yet to be defined.

    - **Mechanisms controlling the access to and usage** of the collected data, given that the data shall be made accessible to the public in a discriminatory-free manner (e.g. micro, small and medium-sized companies, academia and journalism).

    - **Transparency measures** to prevent citizens from feeling to be under constant surveillance, given the scale and potential opacity of the data processing.

- Most importantly, to comply with the principles of data minimisation and purpose limitation (and therefore, supporting the requirements of the "legitimate interests"-clause, as well), a **decentralised system is clearly more suited** to meet these principles since the research purposes do not necessarily require to store the data centrally over a longer period. Instead, it is possible that the system only provides a set of technical and legal standards for how the data must be treated *when* collected and combined.

The third part of the DPIA consists of the **actual risk assessment**. Since the target of evaluation has not yet been sufficiently defined, the risk assessment cannot, as required by Art. 35 GDPR, determine on which carriers the data will be processed, which persons will come into contact with the collected data, the exact threats that may occur and what concrete measures must be taken to reduce the risk. However, the following aspects could be clarified:

- Issues surrounding the **methodology related to the fundamental rights-based approach** of data protection impact assessments. This DPIA is based on the criteria catalogue proposed by the Art. 29 Data Protection Working Party and refers to the risk assessment methods proposed by the French Data Protection Authority (CNIL), and the Standard-Datenschutzmodell (in the following "SDM") by the Technology Working Group ("AK Technik") of the Conference of the German Data Protection Authorities and the Forum Privatheit.

- On this basis, the current DPIA can provide a first assessment on the severity of the impacts caused by the planned system (but not on the likelihood of threats given the target of evaluation as too unspecific). An impact can occur especially on

    - the fundamental right to data protection caused by the data protection risk of the data subjects' **"feeling to be under constant surveillance"**;

    - the fundamental right to private life resulting from the data collection providing insights into the private life of data subjects (especially by **movement patterns**);

    - the right of private autonomy or a fair trial that arise from the risk that the data could be used against the data subjects in relation to **contractual arrangements** (e.g. with an insurance company) or **legal proceedings** (e.g. by the police).

    - the fundamental rights of non-discrimination due to the risk that data subjects are **treated differently** pursuant to their willingness to disclose their data or not (e.g. by the parking lot finder application).

To mitigate the impacts, as to comply not only in the moment of data collection but in particular at a later stage with the requirements of the "legitimate interests"-clause, the implementation of the principles of data minimisation, purpose limitation and transparency, **certificates and codes of conduct** can play a major role. An important question in this case is who will sit on the steering committee (e.g. only representatives of the controller or also of the data subjects).

RECOMMENDATIONS

Based on the findings of this DPIA, the **providers of the envisaged system should define their data protection by design strategy** by conducting a consultation process that involves the data subjects and their representatives. During this consultation process, the above-mentioned additional safeguards should be considered, in particular, in light of the (potential) impact as listed before. Providers of products and/or services that build on the processing of the collected data by the system should make further assessments based on this DPIA – even if it should not be legally required. The reason for this is that these further DPIA's can make a significant contribution to help design their services and products not only in a data protection compliant way but even push the current state of the art (see esp. Art. 25 sect. 1 GDPR). An important question that needs to be addressed in this regard is how the data subjects can effectively deal with the remaining risks and even more so with new data protection risks caused by (previously) unknown purposes. Furthermore, the providers should set up a code of conduct and possibly additional certification schemes to effectively control the risks that particularly result from the later usage of the data. The relevant **steering committees should include representatives** not only of controllers and processors, but equally **of data subjects**.

The **European legislator should reconsider** the approach of its proposal for an ePrivacy Regulation according to which the processing of communications metadata may only be permitted for the purposes of technical transmissions, security, the fulfilment of quality of service requirements — or if it is based on the individual's consent. As it is difficult, if not impossible, to make such metadata completely anonymous, this approach would exclude the processing of such data for research and statistical purposes, for example, in an urban traffic management environment. **Alternative mechanisms, such as an exception for statistical and research purposes (or even for the "legitimate interests" for the controller and the public)**[2] ***if combined with an obligatory*** **adherence to certification schemes or codes of conduct and technical and organisational safeguards**, could ensure an equally effective protection of such data that respects the fundamental right to communication of data subjects, by less restricting the interests of third parties and the public when processing such data.

As far as the GDPR applies, the **German legislators should** consider **specify**ing **the requirements for** the processing of personal data which are necessary for the performance of a task carried out in the public interest (Art. 6 sect. 1 lit. e in combination with sect. 2 and 3 GDPR). The legislators should consider this possibility in particular with respect to the processing of personal data for research and statistical purposes, such as **in an urban traffic management environment**. If this is the case, the legislator should consider the data protection by design requirements as proposed in this DPIA. Furthermore, the **City of Berlin and/or a federal ministry** such as the Ministry of Justice and Consumer Protection, the Ministry for Economy and Energy and/or the Ministry of the Interior, **should actively support all efforts to develop codes of conduct and/or certification schemes**. This could be done financially and/or by assisting the organisation and coordination of the negotiation process between the involved stakeholders (including the competent DPA(s)). The competent DPA(s) should actively support these efforts (e.g. as initiated) in this project by proactively moderating between the potentially conflicting opinions among the German DPAs and between the German and the European level.

[2] See, in contrast, the "Statement of the EDPB on the revision of the ePrivacy Regulation and its impact on the protection of individuals with regard to the privacy and confidentiality of their communications" at https://edpb.europa.eu/our-work-tools/our-documents/drugi/statement-edpb-revision-eprivacy-regulation-and-its-impact_en, last access: 18th May 2020.

CONTENTS

I INTRODUCTION

The introduction of this Data Protection Impact Assessment ("DPIA") will clarify the following aspects: First, why this DPIA is going to be carried out; second, its context, object and approach will be outlined; and third, insights into how this DPIA is designed will be given.

1 Why is a DPIA important / necessary?

The reason for carrying out this DPIA is the likely high risk that is considered to be the result of its evaluation objective. The subject-matter of this DPIA (so far) is a hypothetical smart city application across Berlin. In this hypothetical scenario, CCTV cameras record anonymous data about the flux of pedestrians and vehicles, sensors monitor available parking spaces, and a publically available Wi-Fi system captures data about the use of the devices that communicate with the Wi-Fi network. This data is supposed to be used for (research) purposes in an urban traffic management environment. The DPIA is carried out with a particular view on the General Data Protection Regulation ("GDPR"), which has come into effect in May 2018. Under Art. 35 GDPR, a DPIA is required if the planned processing operations are likely to lead to a high risk for the rights and freedoms of natural persons. This is especially the case, if the data processing consists of a systematic monitoring of publicly accessible areas, Art. 35 sect. 3 lit. c) GDPR. This consideration behind the legal provision serves as an essential reason for the conduct of this DPIA because the planned data collection consists of a systematic monitoring of publicly accessible spaces, expanding to the whole area of the City of Berlin. Regardless of the question, whether the GDPR applies or not, the decision to conduct a DPIA is also based on the assumption that by conducting a DPIA, it can serve as an important means for developing such a smart city application in a socially sustainable way.[3]

In fact, even if the processed data is very likely to be "personal" and therefore falls under the scope of data protection law(s), it does not automatically lead to the application of the GDPR, and its Article 35. For example, with respect to the processing of data in relation to the Wi-Fi system, it is reasonable to consider that this kind of processing will fall under the upcoming ePrivacy Regulation repealing the ePrivacy Directive.[4] Other data protection laws could equally apply. If such data were to be originally collected for (research) purposes in an urban traffic management environment and if this data should be processed, one day, for example, for the prevention, investigation, detection or prosecution of criminal offences, the Directive 2016/680 applies (in the following "Directive for the police and criminal justice sector").[5] Even if the GDPR was applicable, its provisions could be adapted, or at least, specified by means of a legal provision laid down by national law for the performance of a task carried out in the public interest.[6] In all these cases, where Art. 35 GDPR should not apply, the conclusions drawn within this DPIA may nevertheless help to build such a smart city application in a way that respects the principles of data protection and is, therefore, socially sustainable.

[3] Cf. *Friedewald* et al., Forum Privatheit, White Paper Datenschutz-Folgenabschätzung. 2017, p. 20, URL: https://www.forum-privatheit.de/publikationen/white-paper-policy-paper/ ; ISO/IEC 29134:2017 Information technology -- Security techniques -- Guidelines for privacy impact assessment. Last access: 17th March 2020.
[4] Cf. Albrecht/Jotzo, Das neue Datenschutzrecht der EU - Grundlagen, Gesetzgebungsverfahren, Synopse, Nomos, Baden-Baden: Nomos, p. 66 cip. 20, 29
[5] See Art. 2 sect. 2 lit. d) GDPR; *Ernst* in: Paal/Pauly, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 2 cip. 22
[6] See Art. 6 sect. 1 lit. e) in combination with sect. 3 sent. 2 GDPR; *Buchner/Petri* in: Kühling/Buchner, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 6 cip. 114; *Plath* in: Plath BDSG/DSGVO, 3. Aufl., 2018 Art, 6 cip. 16.

2 Context, target and approach of the evaluation

In this section, we will first outline the context of this DPIA, before the DPIA's target and approach will be defined. We will then present an outlook on the legal (and further) implications that this approach has for the participants involved in the future system (i.e. in the object of evaluation). The final section will give an overview of the design for conducting this DPIA.

2.1 Context

Firstly, this chapter describes a technological smart city research project, which is currently conducted in Berlin. This project serves as the essential source of inspiration for the associated interdisciplinary research project "Data protection and security by design in Smart Cities". Secondly, this interdisciplinary project is described because it is the actual research framework for this DPIA.

2.1.1 Smart City Berlin - Ernst Reuter-Platz (Cisco & DAI)

The technological research and development project "Smart City Berlin – Ernst Reuter-Platz" (hereinafter also "technological research project") is funded by "Cisco Systems GmbH" and jointly carried out with the Technische Universität Berlin ("TU Berlin"), the Distributed Artificial Intelligence Laboratory ("DAI-Labor"), and other research partners. This technological research project collected data by monitoring publicly accessible areas in the City of Berlin on a large scale in order to use the data for a variety of applications, which are yet to be specified, in a smart city environment. The collected data stemmed from three main sources:

(1) A publicly accessible Wi-Fi system that recorded the location and time of devices connected to available access points (i.e. monitoring the "people who are carrying these devices".

(2) Camera-based sensors that captured moving objects such as pedestrians and vehicles (i.e. monitoring the "flow of pedestrians and cars").

(3) Camera-based sensors that documented spaces where vehicles are supposed to park (i.e. monitoring whether parking spaces are available or not).

In the framework of the technological research project, the collection of the data (i.e. the system) was restricted to a publicly accessible roundabout, including the sidewalks nearby, the Ernst Reuter-Platz, and adjacent entrances and in the TU Berlin.

The collected data was intended to be anonymized and was primarily supposed to be used for traffic management purposes. The aim of the technological research project was to find out how this data could actually be used, specifically for the, so far, rather generally defined smart city (research) purposes.

The technological research project started in "April 2017" and it was intended to run until the end of 2017. In the event of its success, there may well be support for the idea, amongst the research partners as well as further stakeholders (such as the municipalities and other private companies) of expanding the system throughout the Berlin area. In that case, in order to enhance the innovative capacity of Berlin-based actors in the private sector (e.g. research institutes, companies, journalists, interested communities and citizens), the data should be made publicly available in a non-discriminatory way. However, it was not yet clear, which bodies should carry the specific responsibilities as data controllers and even more so, how to ensure full compliance with the data protection law. For this reason, another research project was set up to address these questions.

2.1.2 Data protection by design in Smart Cities (Cisco & HIIG)

The accompanying research project "Data protection and security by design in Smart Cities (GAMEaTHON)" (in the following hereafter referred to as "privacy by design research project") addresses the following question: how should such a system, which should be implemented on a large scale in real practice, thus, outside of the "sandbox" of a technological research project, be designed to conform to data protection law(s)? This interdisciplinary research project is run by the Alexander von Humboldt Institute for Internet and Society and financed by "Cisco Systems GmbH".

Taking the system developed by the technological research project as a starting point, the current research project examines how such a system should be developed, by applying the required principles of data protection and security by design stated in the GDPR. The main part of this examination is the present DPIA. The DPIA endeavours to evaluate the data protection risks as well as possible implemented technological and organisational measures to avoid, or at least, reduce the risks. The results of the project should ultimately lead to recommendations on how such a system could be entirely controlled.

2.1.3 Resulting challenges for the current DPIA

In view of the described research framework, the DPIA faces the challenge of being only a partially defined target of evaluation. Concerning such a partially defined target of evaluation, the following three questions arise: First, how accurately can and must the target of the current DPIA be defined?

Second, which approach is appropriate to reach the aim of carrying out the DPIA? Lastly, which implications do these results have for the system participants who, one day, may have to act in compliance with the GDPR, in particular, with its Art. 35?

2.2 Target of evaluation, approach and legal implications

The next section first examines in detail whether the target of evaluation of this DPIA can meet the requirements of Art. 35 GDPR. After having denied this question, the next section will illustrate the particular approach of the current DPIA as a combination of a legal and a scientific DPIA. The chapter concludes with the implication that such an approach has for the participants involved in the future system (i.e. in the target of evaluation).

2.2.1 The system as a whole (not only single operations)

Art. 35 sect. 1 GDPR requires that "a single assessment may address a set of similar processing operations that present similar high risks". This raises the question of whether a DPIA can only treat similar risks by excluding other (additional) risks. On the one hand, the legislator appears to have sought, with this wording, to avoid that the scope of assessment will be too general and therefore ignores, neglects or hides specific (relevant) risks.[7] On the other hand, this wording must not lead to the conclusion that a DPIA should not cover a variety of different risks stemming altogether from a common source. In contrast, teleologically, a DPIA should instead cover such a variety of risks that come from a common source because such a broad scope allows the data controller to control the risks caused by its processing activities, comprehensively and consistently.[8] Recital 92 backs this second standpoint stating:

> "[T]there are circumstances under which it may be reasonable and economical for the subject of a data protection impact assessment to be broader than a single project, for example where public authorities or bodies intend to establish a common application or processing platform or where several controllers plan to introduce a common application or processing environment across an industry sector or segment or for a widely used horizontal activity".[9]

Thus, recital 92 makes clear that the current DPIA does not have to be restricted to one single project and even more so, only to similar high risks. Even more so, it can even cover a common processing environment across an industry sector including several controllers or bodies.

This broad scope corresponds to the situation underlying the current DPIA where several bodies collaborate, providing each of them a different piece of the system technology, to establish a common platform for the collection of data from different sources for different processing activities in a Smart City environment. Thus, the system to be established is here principally deemed to be a legitimate target of evaluation, according to Art. 35 GDPR. Such a broad scope of a DPIA is also justified (if not necessary) because it enables the (joint) data controllers to address the overall risks of the system comprehensively. Consequently, such a broad scope also enables the (joint) controllers to implement the

[7] Cf. Art. 35 Sec. 1 GDPR; *Baumgartner*, in: Ehmann/Selmayr, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 35 cip. 2; cf. Friedewald et al., Forum Privatheit, White Paper Datenschutz-Folgenabschätzung. 2017, p. 34, https://www.forum-privatheit.de/publikationen/white-paper-policy-paper/ ; ISO/IEC 29134:2017 Information technology – Security techniques – Guidelines for privacy impact assessment. Last access: 17th March 2020.
[8] See *Martini,* in: Paal/Pauly, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 35 cip. 8 and seq.
[9] Cf. Recital 92 to the GDPR, Broader data protection impact assessment, URL: https://gdpr-info.eu/recitals/no-92/, Last access: 17th March 2020.

necessary technical and organisational protection measures (hereinafter "TOM") consistently.[10]

However, there is another aspect which conflicts with the current DPIA's approach. Under Art. 35 sect. 7 GDPR, it is necessary to specify, at least, "a systematic description of the envisaged processing operations and the purposes of the processing, including, where applicable, the legitimate interest pursued by the controller".[11] Thus, Art. 35 GDPR requires a specific DPIA, which means a target of evaluation that is precisely defined.[12] In light of the circumstances of the current DPIA, this requirement seems to be impossible to meet since the target of assessment is, per se, not (yet) comprehensively defined. As illustrated previously, it is only evident, so far, that the data is supposed to be used in the framework of the technological research project for Smart City (research) purposes. In contrast, it is unclear for which specific purposes the data will be used when the system is running in a real (business) practice covering the entire area of Berlin. Moreover, it is uncertain how to implement this system precisely. In particular, it is not even clear whether the participants of such a future system will be private bodies and/or public bodies, and whether they act as controllers, processors or manufacturers. Insofar, the current DPIA cannot not meet the requirement under Art. 35 sect. 7 GDPR.

2.2.2 Combination of a specific and scientific DPIA

However, since the existence of data protection risks of such a system is obvious, it is useful to carry out a DPIA. In particular, it is helpful to carry out a DPIA already in the conceptual phase. The reason for this is that the impact of the DPIA on a system design is likely higher (and certainly less expensive) than if the TOM were to be defined after one entirely implements the system.[13] This leads to a conflict for the participants of such a system, which are already involved in the conceptual phase: On the one hand, the system per se is not yet sufficiently defined in order to meet the requirement of Art. 35 GDPR; on the other hand, the impact is much higher (and more cost effective) if the DPIA is carried out in such an early stage of the development.

To solve such conflicts, the Forum Privatheit proposes to combine a specific DPIA, as required under Art. 35 GDPR, with a generic, so-called scientific DPIA. The reason for this is that a scientific DPIA addresses the knowledge uncertainties, which are typically arising in an early stage of a technology development, such as the future use of personal data and the corresponding data protection risks.[14] To meet these *uncertainties*, a scientific DPIA uses a prognostic methodology.

One method is, when using a prognostic methodology, to define certain expected, or typical, scenarios about how one may use the technology and/or data.[15] This makes it possible to carry out an assessment on the basis of *specific* knowledge that exists, at least, with respect to these scenarios.[16] In the best case, indeed, one combines both forms of a DPIA. In this case, the DPIA's description consists of two

[10] Cf. Article 29 Data Protection Working Party, Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679, 04 April 2017, 17/EN, WP 248, URL: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611236 p. 7 f., stating: "When the processing operation involves joint controllers, they need to define their respective obligations precisely. Their DPIA should set out which party is responsible for the various measures designed to treat risks and to protect the rights of the data subjects.", last access: 17th March 2020.
[11] Cf. Art. 36 Sec. 7 GDPR, Jandt in Kühling/Buchner, Datenschutz-Grundverordnung, 2. Aufl., 2017, Art. 35 cip. 34-38.
[12] Cf. *Friedewald* et al., Forum Privatheit, White Paper Datenschutz-Folgenabschätzung. 2017, p. 22, URL: https://www.forum-privatheit.de/publikationen/white-paper-policy-paper/; ISO/IEC 29134:2017 Information technology – Security techniques – Guidelines for privacy impact assessment, last access: 17th March 2020.
[13] Cf. Roßnagel, Rechtswissenschaftliche Technikfolgenforschung – Umrisse einer Forschungsdisziplin, Baden-Baden: Nomos, 1993, pp. 182–185.
[14] Cf. *Friedewald* et al., Forum Privatheit, White Paper Datenschutz-Folgenabschätzung. 2017, p. 24
[15] Cf. *Friedewald* et al., Forum Privatheit, White Paper Datenschutz-Folgenabschätzung. 2017, p. 24.
[16] Cf. *Friedewald* et al., Forum Privatheit, White Paper Datenschutz-Folgenabschätzung. 2017, p. 24.

components:

- First, a "systematic description of the envisaged processing operations and the purposes of the processing, including, where applicable, the legitimate interest pursued by the controller(s)", Art. 35 sect. 7 GDPR, will be outlined to the best possible extent.
- Second, the definition of certain data usage scenarios that are particularly suitable in order to evaluate appropriate data protection by design strategies for the system development.[17]

This combined methodology also applies to the current DPIA. The current DPIA will, as a result, carry out the assessment both on the basis of the system as previously described, or more precisely, how it is intended to be implemented in a real business environment across Berlin, and on the basis of three data usage scenarios that are going to be defined in the following.

An additional technique frequently used as part of a prognostic methodology is to involve various stakeholders concerned by the final use of the technology.[18] There are two advantages of such a stakeholder involvement: First, the different perspectives of the stakeholders help to both generate the knowledge about the risks as well as balance potentially conflicting interests. The second advantage is that the technology, which will finally be used in practice, is socially accepted. However, such a stakeholder involvement also has some constraints: For example, such an involvement requires a certain amount of time and effort. Furthermore, concerns regarding business secrets can hamper an involvement of external stakeholders. Often, such an involvement is therefore restricted to the factual circumstances in which a DPIA is carried out.[19]

Nonetheless, the current DPIA involves external stakeholders in order to gather further necessary knowledge and to represent the conflicting interests of most of the involved stakeholders. However, in light of the restricted time frame and budget of the data protection by design research project, the stakeholder involvement will follow a "lean" approach.

2.2.3 Legal and further implications

The legal implication of this (necessarily) methodological combination is that the current DPIA does not accomplish the requirement stated by Art. 35 GDPR. However, there are three significant advantages of carrying out the current DPIA:

- The results of the current DPIA provide the basis for the internal development of a data protection by design strategy. This strategy helps the current system developers in a way that finally meets the data protection requirements under Art. 35 Sec. 7 GDPR.[20]

- The results also provide a vital knowledge basis for external participants, which are not yet involved in the current system development but will later use the technology and/or the collected data for new (own) purposes.

    - Based on this information, i.e. about common, or infrastructural data protection risks, the (eventual) data controller can then be guided, first of all, how to use the technology responsibly.

[17] Cf. *Friedewald* et al., Forum Privatheit, White Paper Datenschutz-Folgenabschätzung. 2017, pp. 24.
[18] See also Art. 35 sect. 9 GDPR.
[19] Cf. *Friedewald* et al., Forum Privatheit, White Paper Datenschutz-Folgenabschätzung. 2017, pp. 29-30.
[20] Cf. *Friedewald* et al., Forum Privatheit, White Paper Datenschutz-Folgenabschätzung. 2017, pp. 22-23.

- Furthermore, the results of the current DPIA will also provide an essential basis for specific DPIAs that may be necessary later on, under Art. 35 GDPR.[21] In this case, the pre-assessment carried out by the current DPIA significantly helps to open "the way for scalability, meaning that even a small data controller can design and implement a suitable DPIA", by building upon the current one.[22]

- Finally, and maybe most importantly, the third advantage is that the current DPIA, (at a very early stage), signals to all interested parties that their concerns are indeed playing a crucial role in the development of the system. The current DPIA, therefore, helps build trust amongst Berlin citizens (and all other people being there) in the proposed system. This will be particularly the case when the current DPIA is publicly accessible.[23]

2.3 Design of the DPIA

This chapter specifies the design of this DPIA, which means: the guidelines used to conduct the assessment; the timeframe and contractual arrangements; as well as the people involved in this DPIA.

2.3.1 Guidelines

The current DPIA follows the guidelines proposed by the Art. 29 Data Protection Working Group on DPIAs. Referring to Art. 35 sect. 7 of the GDPR, as well as its recitals 84 and 90, the Working Group highlights that the GDPR provides "a broad, generic framework for designing and carrying out a DPIA".[24] Such a broad framework is supposed to provide "data controllers with [the] flexibility to determine the precise structure and form of the DPIA in order to allow for this to fit with existing working practices."[25] The Working Group has finally set up common criteria to allow different methodologies for carrying out a DPIA, "whilst allowing controllers to comply with the GDPR".[26] As a result, the current DPIA is primarily found on these common criteria.

However, there are also further guidelines to which the current DPIA refers. As already mentioned earlier, the "White Paper, Datenschutz-Folgenabschätzung: Ein Werkzeug für einen besseren Datenschutz" by the Forum Privatheit serves as an important source. Also, the "PIA, Methodology" by the French Commission Nationale de l'Informatique et des Libertés" (in the following "CNIL"), is an important source, not least because of the templates describing the data protection risks.[27] Finally, the

[21] Article 29 Data Protection Working Party, Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679, 04 April 2017, 17/EN, WP 248, p. 7, URL: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611236

[22] See Article 29 Data Protection Working Party, Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679, 04 April 2017, 17/EN, WP 248, p. 15, URL: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611236, last access: 17th March 2020.

[23] See Article 29 Data Protection Working Party, Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679, 04 April 2017, 17/EN, WP 248, p. 17; Forum Privatheit: White Paper, Datenschutz-Folgenabschätzung: Ein Werkzeug für einen besseren Datenschutz, 2016, pp. 21, URL: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611236, last access: 17th March 2020.

[24] See Article 29 Data Protection Working Party, Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679, 04 April 2017, 17/EN, WP 248, pp. 14 and 15, URL: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611236, last access: 17th March 2020.

[25] See Article 29 Data Protection Working Party, Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679, 04 April 2017, 17/EN, WP 248, p. 15, URL: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611236, last access: 17th March 2020.

[26] See Article 29 Data Protection Working Party, Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679, 04 April 2017, 17/EN, WP 248, p. 15, URL: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611236, last access: 17th March 2020.

[27] See the internet knowledge base "PIA, Tools", publically available under https://www.cnil.fr/sites/default/files/typo/document/CNIL-PIA-2-Tools.pdf, last access: 17th March 2020.

ISO standard "Information technology — Security techniques — Privacy impact assessment – Guidelines" ISO/IEC FDIS 29134:2017(E) also serves as an essential basis for the structure of this current document.

2.3.2 Timeframe and contractual arrangement(s)

As mentioned in the above, the data protection by design research project is conducted by the Humboldt Institute and financed by "Cisco Systems GmbH". The research project already started in April 2017 and will run until around the first quarter of 2018. Whilst conducting the data protection by design research project, the Humboldt Institute is collaborating with subcontractors to fulfil specific tasks. One of the sub-contractors is the Berlin-based law firm iRights.Law, which is responsible for implementing the current DPIA.

After an informal preparation phase beginning in June 2017, the first drafting of the DPIA has started in August 2017 and was finished in the end of 2017. The conduct of the DPIA has been structured along the following points:

Aug to Nov 2017	First Draft of the DPIA;
Nov 2018	Workshop based on the first draft of the current DPIA;
Feb 2020	Second, final draft of the current DPIA, based on the feedback from the workshop and additional research.

2.3.3 Involved people

This DPIA partly builds on a study assessing the legal aspects of the target evaluation, commissioned to the law firm iRights.Law, as well as on a short assessment by the academic spin-off Law & Innovation on the role of certification schemes and codes of conduct. During the preparation phase and whilst compiling the first and second draft of this DPIA, informal information and feedback was, in particular, given by:[28]

Klaus Lenssen (Chief Security Officer - Head of Security & Trust Office Germany, Cisco)
Uwe Northmann (Sales Business Development Manager, Cisco Germany)
Manuel Beicht (Solutions Integration Architect, Cisco Germany)
Martin Berger (Researcher - CC Education, DAI-Labor)
Sebastian Berg (Researcher - DAI-Labor)

During the workshop, feedback was given by:

Jörg Pohle (Researcher - HIIG)
Kevin Klug (Research assistant - HIIG)
Christopher Olk (Research assistant - HIIG)
Klaus Lenssen (Chief Security Officer - Head of Security & Trust Office Germany, Cisco)
Claus Schaale (Cloud Leader - Datacenter EMEAR Cisco)
Staff members of the Berlin data protection authority[29]

[28] Cp. Article 29 Data Protection Working Party, Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679, 04 April 2017, 17/EN, WP 248, p. 13 and 14, https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611236, last access: 17th March 2020.
[29] It is important to note again that we asked the DPA only for their expert feedback, not for an official statement; therefore, the

3 Description of the target of evaluation

This DPIA continues with the description of its target of evaluation. Following the context of this DPIA, the next sections will begin by describing the data collection framework currently being set up as part of the technological Smart City research project. With the approach of combining a legal and a scientific assessment, this DPIA then defines three scenarios of how the data might be used when implementing the system in real business practice across Berlin.

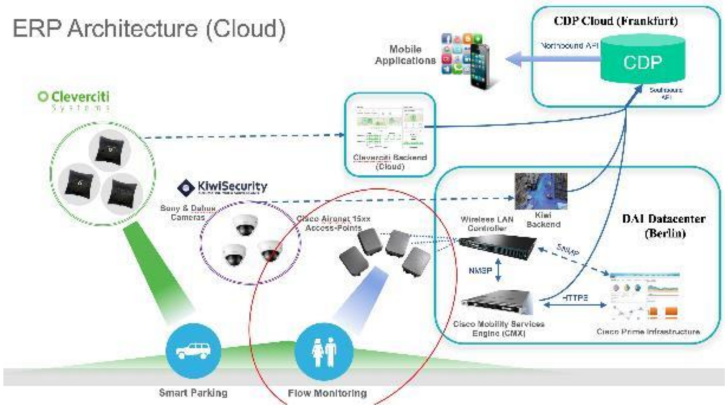
3.1 Data collection framework

As already mentioned, in the current technological research project, there are three data collection sources: First, the publicly accessible Wi-Fi system provided by Cisco Germany; second, the CCTV cameras provided by KiwiSecurity; and third, the sensors capturing the parking lots, which are provided by Cleverciti. The following sections first describe how (and where) the data is currently collected, processed, and stored. The last section gives a brief summary highlighting the most relevant aspects (of the system described), which may either be upheld, when the system will be implemented in real business practices across Berlin or changed and/or further specified.

3.1.1 Current data collection (and protection) framework

The following picture illustrates where the collected data is processed and stored: With regard to the parking space sensors, the collected data is directly processed within the sensors itself; after the first processing, the data is transferred to Cleverciti's backend and, from there, transferred to Cisco's Smart City Cloud. Cisco's Smart City Cloud is a data storage centre located in Frankfurt, Germany (in the following "CDP Cloud"). However, the transfer of the collected data via the CCTV cameras and the Wi-Fi access points is more complex: This kind of data is first transferred to virtual machines (in the following "VM") by KiwiSecurity and Cisco, where the collected data is processed or "anonymized"); then, the data is transferred to the DAI Datacenter in Berlin for further processing research purposes; finally, this processed data is transferred back, to the CDP Cloud where it is stored for further (yet unspecified) purposes.

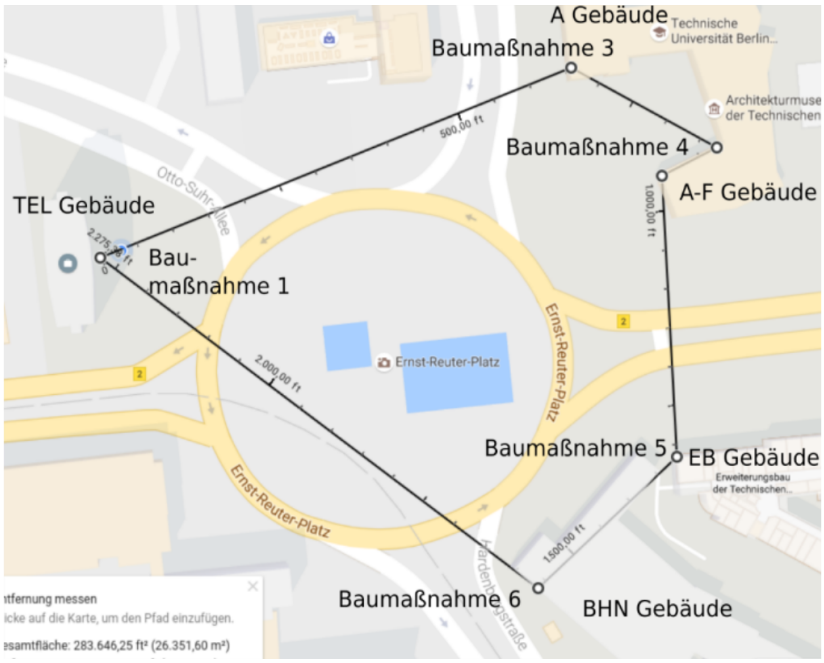
Overview



The next picture emphasises how the data is currently protected against an intrusion from outside of the systems: All the data is collected through three virtual subnetworks, which are interconnected and accessible from outside via an umbrella of virtual private networks. The virtual private networks are

final findings and recommendations of this DPIA should be understood as our own position and not that of the DPA.

provided by the technological research partner TU Berlin. Each subnetwork covers another area of the ERP and consists of CCTV cameras, parking lot sensors and access points for the Wi-Fi network. The first subnetwork includes these "data collectors" installed at the buildings A and A-F (in the above right corner of the picture), the second subnetwork consists of the "collectors" installed at buildings EB and BHN (in the below right corner of the picture). The third subnetwork does not only consist of the previously mentioned collectors but also includes VM that are provided by KiwiSecurity and Cisco (on the left side of the picture). The next-following sections will describe, in more detail, how the data is collected and explicitly processed by these VM.



3.1.2 Wi-Fi data collected ("anonymised")

With respect to the Wi-Fi system, this system collects data at the moment where a portable device, which consists of a switched-on Wi-Fi client, gets into the reach of an access point of this Wi-Fi system. The data that are collected basically are:

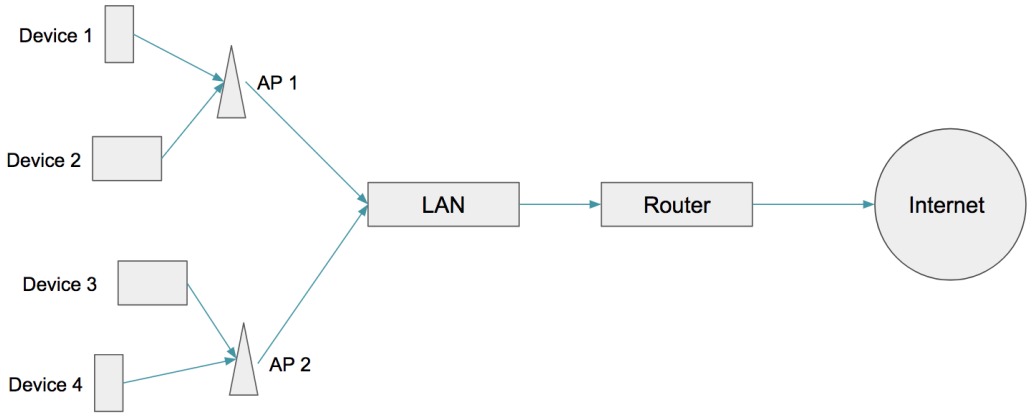
- the Media Access Control ("MAC") address;
- the time and location where these identifiers were/are captured.[30]

This information is renewed every 30 to 60 seconds, i.e. the system collects every 30 to 60 seconds the time and location of the IMEI and MAC address(es) captured. In order to understand the collection of these types of data, it is necessary to illustrate how such Wi-Fi systems work.

Wi-Fi systems typically operate in a so-called infrastructure mode to grant access to the world wide web (whereas the so-called ad-hoc mode can be used for simply enabling data exchange between devices). The infrastructure mode implies that data is not directly transferred between devices but always uses a

[30] In fact, more data is collected, such as the Information Elements of probe requests (outlined in the subsequent description of Wi-Fi systems) and sequential numbers see *Eisele* et al., Forum Privatheit, White Paper Privatheit in öffentlichen WLANs, 2017, p. 34, referring to *Vanhoef* et al., Why MAC Address Randomization is not Enough: An Analysis of Wi-Fi Network Discovery Mechanisms. In: Proceedings of the 11th ACS on Asia Conference on Computer and Communications Security. ACM, 2014, p. 413-424.).

third central entity (the "Wi-Fi access provider"). Usually, a device containing a Wi-Fi client (which must be turned on) connects to one of several access points (AP) spread over a specific or particular local space. The AP coordinates the connection of all devices and establishes, in turn, the link to the internet (via the LAN and the router).[31]



There are two mechanisms to connect a device and an AP. Under the passive discovery mechanism, the AP sends, at frequent intervals, so-called signal-frames, which mainly contain its MAC address and network name (e.g. "HIIG"). Each time when a device receives this signal, the device displays on its screen the network availability of the AP to the owner of the device. In contrast, under the active discovery mechanism, the device itself "looks for" available APs by sending so-called probe requests to all potentially available APs. These probe requests consist of data about, first, the device's MAC address, second, technical details, and the names of networks to which the device has previously been connected to. The reason for sending these network names is that it can make the connection much more feasible for the user of the device. The idea behind this thought is that a user can often only connect to the internet via an AP if he or she authenticates and/or decrypts the data transfer by typing in a certain password. If a device recognises an AP (or the network) where its user has already authenticated and/or decrypted the connection before, a new authentication and/or decryption is not necessary anymore. In any case, at the moment the AP has established the connection with the device, the AP attributes a certain IP address to the device. This IP address is necessary for machines (and humans) to find and communicate with each other (humans usually use domain names, for example, www.hiig.de, which represent the IP addresses).[32]

Coming back to the technical situation of the Wi-Fi system installed as part of the technological research project, the probe request data is sent in real time from the access point(s) to the Wireless LAN Controller, i.e. the first VM, and from there to the second VM, i.e. the so-called Cisco Mobility Service Engine ("CMX"). Both VMs are, as mentioned previously, part of the VPN hosted by the TU Berlin. In the CMX, the data can be processed for different purposes. In particular, MAC addresses may be hashed before being sent to a third party. However, the hashing of MAC addresses is optional. Furthermore, there is a list of MAC addresses, which is always stored within the CMX in order to allow the recognition of "repeating visitors". This list cannot be read out by another entity other than the VM itself. Thus, there is neither a Graphical User Interface ("GUI") nor a Call Level Interface ("CLI")

[31] See *Eisele et al.*, Forum Privatheit , White Paper Privatheit in öffentlichen WLANs, 2017, pp. 16-18. URL: https://www.forum-privatheit.de/publikationen/white-paper-policy-paper/, last access: 17th March 2020.
[32] See *Eisele* et al., Forum Privatheit, White Paper Privatheit in öffentlichen WLANs, 2017, pp. 18-21. URL: https://www.forum-privatheit.de/publikationen/white-paper-policy-paper/, last access: 17th March 2020.

serving as a standard Application Programming Interface ("API"), which could grant access to this list.

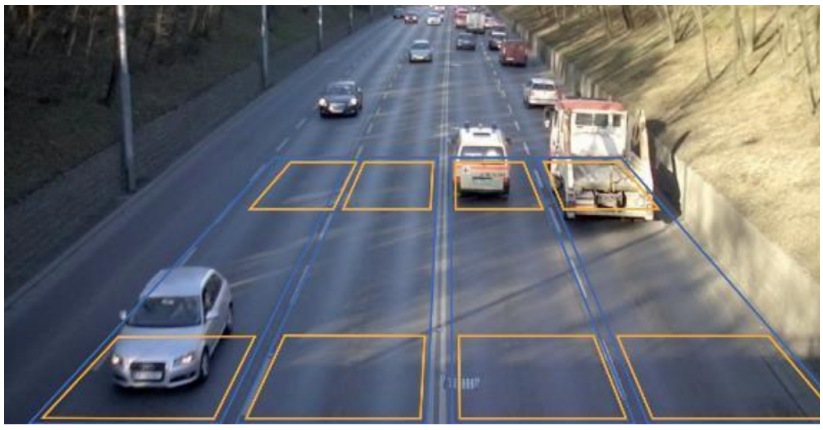
3.1.3 CCTV camera data collected ("anonymised")

Regarding the data collected via the CCTV cameras, the cameras principally collect data if and as long as these cameras are switched on. The technology used by the provider KiwiSecurity offers a large number of possibilities of primary purposes for which the recorded visual data can be analyzed. For instance, it is possible to:

- analyse the direction of objects moving through a captured space (e.g. vehicles or persons);
- count the number of objects moving through such a space;
- and, based on these results to analyse the patterns of traffic evolving over specific periods of time.

In terms of granularity, it is also possible to analyse which kind of objects appear within a captured space, as for example the different types of cars. For the current DPIA, it is important to highlight the following functions that are not used for the current technological research project but that are, in principle, possible:[33]

- license plates of vehicles could be captured and analysed (e.g. comparing them with license plate numbers in a "black or white list" to trace certain cars);
- and, comparably, the faces of data subjects could also be recognised (and consequently be traced).



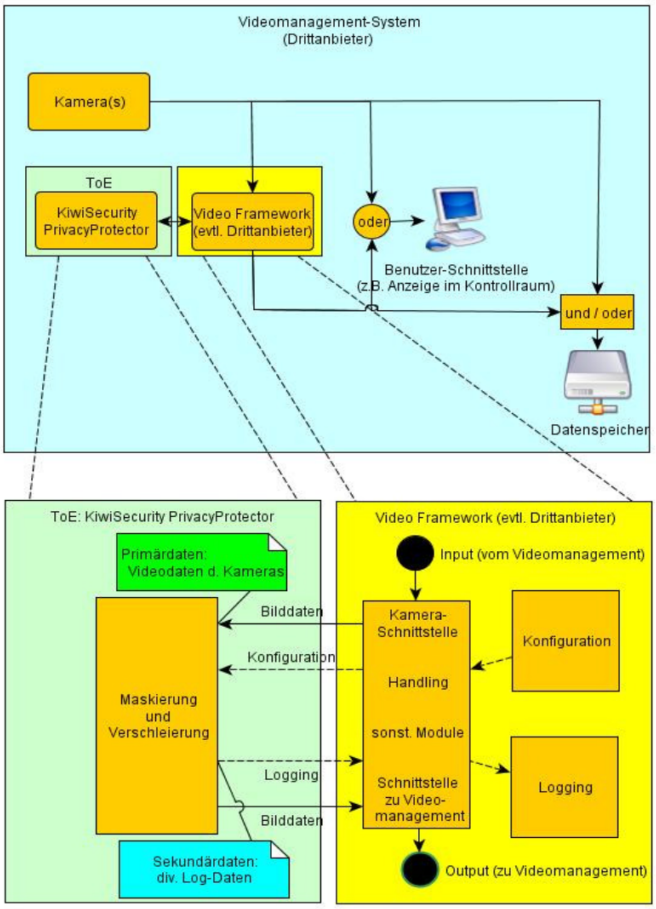
[33] For a more detailed illustration of the use cases, see the following site: https://www.kiwisecurity.com/wp-content/uploads/2016/02/Produktblatt-KiwiVision-Traffic-Analyzer.pdf., last access: 17th March 2020.

However, for the current technological research project, the provider of the CCTV cameras (i.e. KiwiSecurity) uses a privacy-protecting module that is called PrivacyProtector. This module obfuscates certain visual areas of the recorded pictures (e.g. license plates of vehicles and/or human faces) so that the recorded data does not identify an individual anymore.[34]

At the time of conducting this DPIA in 2017, the visual data was immediately encrypted (i.e. within the camera itself) after the recording and sent to the back-end server of KiwiSecurity (to the VM of KiwiSecurity) using a Secure Sockets Layer ("SSL").[35] As mentioned previously, both the back-end server and the VM are part of the VPN hosted by the TU Berlin. At this stage, KiwiSecurity describes the data as "anonymised". Particularly focusing on faces of data subjects, the privacy certification report of the "European Privacy Seal" describes the framework surrounding the "anonymisation" process and the process itself as:[36]

> "KiwiSecurity's Privacy Protector is a software module for integration in a video management system. The Privacy Protector can be called and configured by a video-framework (delivered by KiwiSecurity or third parties). The framework receives video data from surveillance cameras and hands these data over to the Privacy Protector module for obfuscation. The module analyses the incoming video data, recognises persons standing or moving within the scene and obfuscates them. The obfuscated video data is then passed to third party software systems for display and / or storage via the video framework. Different obfuscation mechanisms are available. These can be selected and configured via the video framework, which also continuously takes back log data from the Privacy Protector and stores them into log files or a database. These log data do not contain any personally identifiable data. Configuration of KiwiSecurity's Privacy Protector is carried out by KiwiSecurity or personnel specially trained and certified by KiwiSecurity. (...) In practice, the Privacy Protector will operate between the video signal of surveillance cameras and the picture displayed on the monitor or the storage device. Furthermore it can be used to obfuscate video data retrieved from a storage device. This results in displaying just obfuscated video data, with the monitored persons not identifiable."

[34] See https://www.kiwisecurity.com/wp-content/uploads/2016/02/Produktblatt-KiwiVision-Privacy-Protector.pdf., last access: 17th March 2020.
[35] Unfortunately the original website, which has been accessed in 2017 as a reference for this statement (originally, at https://www.kiwisecurity.com/connection-platform/?lang=de) is not accessible anymore. Last attempted access: 17th March 2020.
[36] See the Short Public Report for the recertification of the module in 2015 at https://www.european-privacy-seal.eu/EPS-en/privacy-protector, p. 2.



The "anonymisation" is technically irreversible. However, the original raw data, which still allows the identification of data subjects (e.g. by showing the faces of data subjects), is not automatically deleted immediately after it has been anonymised. Instead, it is possible to choose whether to delete or store the data. In the second case, "access to the non-obfuscated video data can be reserved to specially authorised users (supervisors etc.)."[37]

3.1.4 Parking lot data collected ("non-personal")

Finally, the parking lot sensors scan certain spaces (all ranging in size depending on the particular sensor technology that is installed) by focusing on specific areas, which have to be specified beforehand (how these areas can be specified, will be explained subsequently). Every three seconds the recorded data is updated (i.e. the scan is repeated) and stored as well as processed in the sensors itself. This data contains the following information:[38]

- the size of a parking space in the area that was previously defined;
- the GPS-position of a parking spot;
- and, whether there is a particular object on this parking lot (i.e. whether or not a vehicle has parked

[37] See the following site at the end of point 7 of the Short Public Report for the recertification of the module in 2015 at https://www.european-privacy-seal.eu/EPS-en/privacy-protector, p. 3.
[38] See https://www.cleverciti.com/wp-content/uploads/2020/01/cleverciti_Sensor_1page_WEB.pdf, last access: 17th March 2020.

there).

As mentioned in the above, the originally scanned raw data does not leave the sensors. Only the processing results are transferred in real-time, after having been encrypted through WLAN and/or 3G from the sensors to the back-end of the provider (i.e. Cleverciti).[39] Here, the transferred data is either stored in the cloud or on a server of Cleverciti. The cloud and the server are not part of the VPN hosted by the TU Berlin. Third parties, i.e. customers of Cleverciti, obtain access to that data via an API. Through this API, the data can also be combined with further data, for example, with data from a ticket machine indicating if the user of a parking space has paid the parking fee or not. To use the analytical results, the provider offers its customers the possibility to analyse the data further and visualise the results via a so-called cockpit. Here, the customer can use the following functions:[40]

- To obtain access to live control and occupancy data;
- to generate additional revenues by optimising the usage of the space and by adapting the pricing dynamically;
- to reduce the cost of enforcement by prioritising parking areas with a high rate of violations.

As a precondition for these functions, it is necessary to define the parking spaces beforehand. In this regard, the provider offers its customer another user interface, the so-called dashboard. In the dashboard, the user has the following functions:[41]

- Defining or changing parking areas;
- identifying special or restricted areas and setting the correct pricing at any time of the day, or blocking certain parking spaces or zones for a specific time (because of construction work for example)

[39] See https://www.cleverciti.com/wp-content/uploads/2020/01/cleverciti_Sensor_1page_WEB.pdf, last access: 17th March 2020.
[40] See https://www.cleverciti.com/wp-content/uploads/2020/01/cleverciti_Sensor_1page_WEB.pdf, last access: 17th March 2020.
[41] See https://www.cleverciti.com/wp-content/uploads/2020/01/cleverciti_Sensor_1page_WEB.pdf, last access: 17th March 2020.

3.1.5 Outlook on the implementation in a business environment across Berlin (including further system participants)

As mentioned previously, the system cannot yet be comprehensively described. So far the previous descriptions can only serve as an indicator of how the future system could look like. However, it is possible to draw a first conclusion on the aspects that may be particularly relevant to design the future system in a way that could be compliant with the GDPR.

One question that must certainly be clarified is, who the controller(s) of the processing operations will be? Under Art. 4 sect. 7 GDPR, the term "controller' means the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data".[42] In contrast, the processor is "a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller", Art. 4 sect. 8 GDPR.[43] So far, it is mainly the DAI-Labor that is likely to be considered the controller because it sets the research purposes, as (more or less) defined above, for which most data is collected. Whereas, the providers of the Wi-Fi system (i.e. Cisco) and the CCTV cameras (i.e. KiwiSecurity) seem to collect and process the data on behalf of the DAI-Labor (as the controller). Finally, the situation is less clear with regards to Cleverciti that provides the parking lot sensors. The reason for this is that the corresponding data is not transferred directly to the data centre of the DAI-Labor, where it is processed for the research purposes. Instead it is sent (more directly) to the Smart City Cloud. Thus, it may be plausible to argue that Cleverciti has to be considered as setting the processing purposes and act as the controller. Another related question is for which purpose(s) the data is actually stored within the Smart City Cloud, in which all the data currently is stored. However, it is not the aim of the current DPIA to evaluate the present system but the future system. In regard to the future system, this DPIA sets out to give useful recommendations on how to define the future processing purposes and aims to identify the controller(s) that most effectively control(s) the data protection risks.

Another question that has to be clarified is how to control the different conditions under which the participants can access and use the collected data. In the current technological research project, the TU

[42] Cf. Art. 4 Sec. 7 GDPR, *Schild*, in: BeckOK DatenschutzR, 31. Ed., 2019, Art. 4 cip. 87-93c; *Ernst* in: Paal/Pauly, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 4 cip. 55.
[43] Cf. Art. 4 Sec. 8 GDPR, *Schreiber*, in: *Plath* in: Plath BDSG/DSGVO, 3. Aufl., 2018, Art, 4 cip. 28; European Union Agency for Fundamental Rights, Handbook on European data protection law, 2014, URL: http://fra.europa.eu/en/publication/2014/handbook-european-data-protection-law, p. 52-54.

Berlin's VPN protects a main part of the system against illegitimate access from outside. However, how will that access be controlled if the TU Berlin does not participate in the system anymore, but several private companies instead? Moreover, who will guarantee that this data is really anonymised or, at least, pseudonymised and that it will also be so in the future? For example, who will control that the visual data collected by the CCTV cameras is not used to scan faces and license plates, contrary to the present situation? Even if the data is considered to be "anonymised", how can the situation be controlled in which third parties are able (and willing) to refer it largely to an individual, at a later stage? This last question becomes particularly relevant when considering possible scenarios regarding the installed parking space sensors. One function of this service is to monitor whether vehicles are parking in a restricted area or not. Another function is to help to enforce such parking violations. Even if this data is *per se* "non-personal", it is clear that an enforcement only makes sense if an "enforcing entity" shows up on-site and identifies the violating individual. In the described Smart City scenario at large, it is impossible to comprehensively determine in advance what may be possible, if the different data sets are combined with each other and with data coming from outside the system (particularly when the purposes are unknown). To understand, at least on a basic level, the risks resulting from such a future system, and to draw conclusions for potential data protection by design strategies, this DPIA can serve as a very useful starting base.

3.2 Data usage scenarios

To understand potential risks of future processing activities better, even though the future system design is still unclear, the following sections will determine three scenarios to identify how the data may be used. The first scenario refers to the data usage for the purpose of urban traffic management; the second scenario covers the situation where Cleverciti uses the data for the provision of its parking lot assistant. The third data usage scenario concerns the usage of the data for purposes of traffic law enforcement.

3.2.1 Urban traffic management (no identification intended)

In the first data usage scenario the collected data is anonymised or, at least, pseudonymised as far as possible and processed for "urban traffic management purposes". One example for such a purpose is the data that is processed to predict better how many people need certain means of transport at a certain place at a certain time. The aim of such calculations is to plan and organise traffic more efficiently. For example:

- The respective schedules;
- the location of the stops and driving routes;
- and/or where vehicles are parked.

This urban traffic management approach can be applied to all sorts of transports, such as public buses and trains or private bike and car-sharing services. Ideally, there could be a service that provides the user with the "best"[44] route constantly adjusted, regardless of whether and where the user spontaneously changes, his or her previously suggested route.

In order to come close to such an ideal case, the current DPIA assumes that all the collected data is intended to be combined and processed. In regard to the data that is collected via the CCTV cameras,

[44] In this case, the term "best" takes the meaning of "adapting to the personal preferences of the user", such as "fastest", "cheapest" or "most ecological route" et cetera.

this means that this type of data is processed, as illustrated previously, in order to:

- Analyse the direction of objects moving through a captured space (e.g. vehicles and persons);
- count the number of objects moving through such a space in specific timeframes;
- based on these results, analyse the patterns of traffic at certain places that are evolving over certain periods of time.

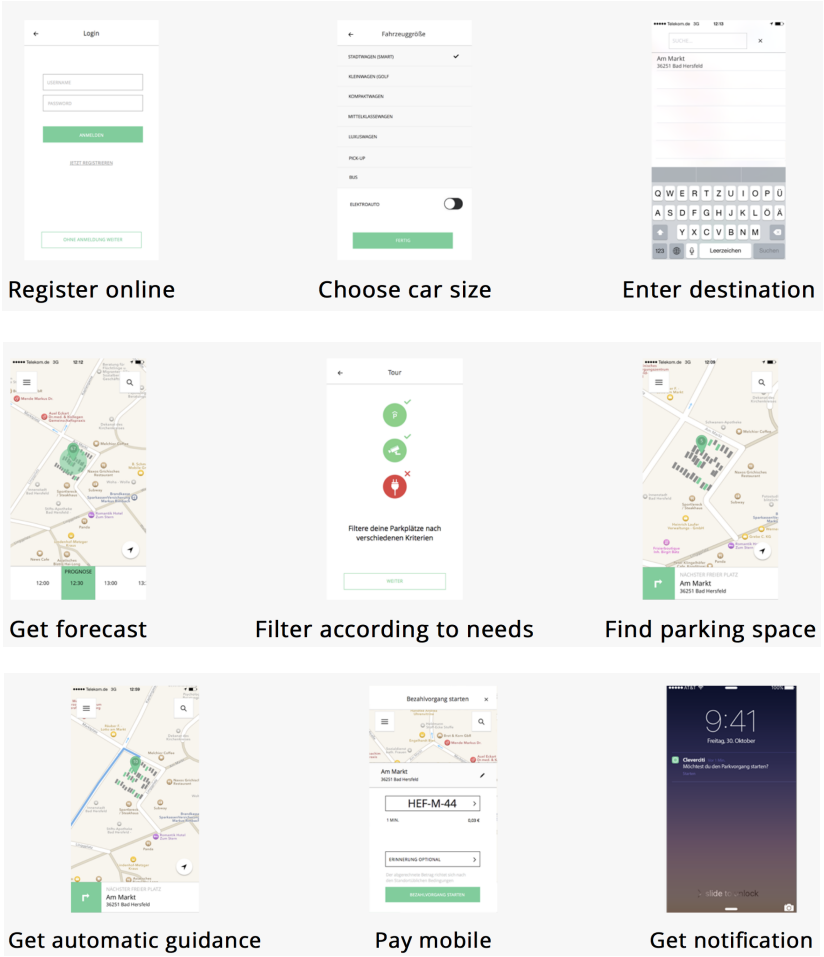
This data is combined with the data from the Wi-Fi system containing information about the time and location of each device (e.g. through the MAC addresses), which is updated every 30 to 60 seconds, and thus allowing rather precise movement patterns, not only about "anonymous masses" but also of each single individual (i.e. owner and/or holder of such a device). Finally, this data is then again combined with the information that is gathered through the parking lot sensors (i.e. where at certain times parking lots are occupied or not). By doing so, traffic planners may be better able to organise where more or less parking lots are needed. In principle, it is also possible to understand which environmental conditions lead, or at least correspond, to an increase or decrease of traffic law violations. This makes it also possible to better plan urban traffic in a way that "pushes" citizens less into conflict with the law. This may be the case, for example, as the City can immediately provide for more public parking when needed, so people "do not have" to park in the second lane.

Of course, similar traffic analysis already exists but with less accuracy. Conventional traffic census is usually done by simply counting the number of vehicles at certain hot lots.[45] In addition, they can be augmented with empirical research that is based on qualitative measurements (e.g. interviews), which can also provide further insights into movement patterns and similar results.[46] However, these traditional methods are limited by their usual errors and shortcomings, such as low depth, unclear reliability, problems with the representativeness of the sample and relatively high costs.[47] Compared to such traditional methods, the approach outlined in the above could bring significant advantages and promises a much better but also different quality of traffic planning. However, it is not the task of this DPIA to assess whether the potential of this approach is really true or not. Rather this usage scenario takes the potential as a given and then asks whether the data processing for this research question could comply with the data protection law or not.

3.2.2 Parking lot assistant (based on user's consent)

The second data usage scenario builds on the previous scenario. However, the user of a personal device now connects to the public Wi-Fi system to use a personalised service. One example for such a personalised service is the parking lot assistant offered by Cleverciti. This service provider explained on its website in 2017 how its service works and illustrates the different stages of a user's process when using this service:

[45] Cf. https://www.berlin.de/senuvk/verkehr/lenkung/vlb/download/Ergebnisbericht_2014.pdf for 2014 results of Berlin.
[46] See for example, https://tu-dresden.de/bu/verkehr/ivs/import/vip/srv.
[47] Cf. for example, *Härtler G.*, Statistisch gesichert und trotzdem falsch?: Vom (Un-)Wesen statistischer Schlüsse. Springer Berlin Heidelberg, 2014; Atteslander, P., Methoden der empirischen Sozialforschung, Erich Schmidt Verlag, Berlin, 2010, pp. 102-103.



Another functionality that is not shown on these screenshots is that a user can also reserve a specific parking space. In this case, the system technically processes the same data, but the impact on data subjects can be much stronger. This is not only the case if such a reservation requires users to pay, i.e. additional payment data. Rather, such a functionality also bears the risk of discriminating against others who are, for example, unwilling to disclose their data or pay for it. In such a case, these data subjects may be harmed because paying users might always be prioritised when searching for a parking space. Indeed, prioritising somebody because they pay for something is an old principle, if not an essential precondition for a market economy. Whether or not such discriminations are justified is therefore a normative decision that must be negotiated in regard to the conflicting values of society. However, the example should make it clear that this second data usage scenario can at least potentially lead to social conflicts.

3.2.3 Law enforcement (identification intended)

In the third data usage scenario, the collected data will be used for law enforcement purposes. According to the first alternative of this scenario, the public regulatory authority gains access to the collected data that is provided by the parking lot sensors provider (i.e. Cleverciti) and, per se, non-personal. Through the cockpit, the regulatory authority can see where (and how close) the vehicles park (to the road). So far, this data might be considered as non-personal because an individual cannot directly be identified due to this type of data per se (in particular, this data does not capture the license plate). In a second step, however, the public order agency uses the so-called dashboard of the service. This additional interface enables the public order agency to identify certain areas as restricted areas according to information

provided by the local traffic street regulation (e.g. stopping restrictions). Thus, the public order agency could observe in real time, if and where vehicles stop in prohibited zones and if a vehicle driver commits an administrative offence against such road traffic law. This real time information allows the public order agency therefore to instantly send their security staff to the "crime scene" to prove the situation. On-site it is now possible to note the license plate number of the respective vehicle and thus to refer the previously non-personal data to an individual, i.e. the owner of the vehicle.

In the second alternative, the data processing activities are more complex. The collected "anonymised" data by the CCTV cameras installed by KiwiSecurity are combined with the collected data through access points of the public Wi-Fi system. As a first step, a public order agency will gain access to the "anonymously" collected data through the CCTV cameras to monitor if "anonymous" persons are crossing a street where they are not allowed to, under the applicable road traffic law. As soon as the public order agency discovers such a "delinquent" (yet "anonymous") person, the agency uses their access to the collected data via the Wi-Fi system in a second step. This enables them to find out which mobile device that is connected to the Wi-Fi system matches with this delinquent person. In the case of a match, the public order agency can now identify, at least in our hypothetical case, who that person is. There are two main presumptions for this result: Firstly, the MAC address of the device collected by the Wi-Fi system is not previously hashed. Thus, the MAC address serves as a unique identifier.[48] Secondly, the agency may also connect the MAC address to the current owner of the device. This requires several further pre-conditions. First, there must be at least one entity that can make the connection between the MAC address and an individual using the device. This may be the internet access provider or just an online service provider. Second, the device must not have changed hands. This means that the original owner still owns the device (and has not given it away, before the administrative offence). If all these requirements are met the "delinquent" person can now be identified. Last but not least, one side note should be made: because the public agency proves this law enforcement mechanism so effectively, the whole process is automatized. As a result, this procedure is now also cost efficient.[49]

3.3 Further expected (but yet unspecified) scenarios

In fact, these scenarios are only few among the imaginable possibilities. In addition to the second scenario, there are many more online services that could be offered in the market, based on the collected data as described. In addition to the third scenario, there are many more data-driven applications ranging from the resolution of legal claims, such as in the insurance sector, to the prevention and prosecution of crimes.[50] Who was where and when are answers to questions of unimaginable value. From a societal point of view however, it is this unimaginable value that makes the collection of the data so critical. If the system is to conform to our idea of a democratic civil liberty society, the collection and use of this data must be monitored in some way to avoid that we run into a total surveillance society. Of course, the question is if such a system can be sufficiently controlled at all.

[48] See Article 29 Data Protection Working Party, Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679, 04 April 2017, 17/EN, WP 248, pp. 8, 13.
[49] According to participants of the technological research project, such a scenario has actually been thought through and would technically be possible.
[50] Cf. *Friedewald* et al., Forum Privatheit, White Paper Datenschutz-Folgenabschätzung. 2017, pp. 22–26.

II CONTROLS FOR URBAN TRAFFIC MANAGEMENT SCENARIO

Data protection laws aim to control the risks resulting from the automated processing of data that is related to data subjects. These laws do not only address risks for data subjects but also concern the whole society.[51] This is a rather broad approach, which can make a risk assessment rather complex. Therefore, to reduce the complexity of the assessment, this DPIA focuses in its legal compliance part on the first described data usage scenario, i.e. data processing for research and statistical purposes in an urban traffic management environment. This does not mean that the other data usage scenarios are not taken into account at all. Instead, there are several entry points within this assessment where the other usage scenarios are very useful to clarify certain legal questions.

1 Application of data protection laws

When assessing the legal compliance with data protection laws, the first question is whether data protection laws apply, which means, whether the collected data relates to an (at least) identifiable individual or not. This first step of the assessment will illustrate that this legal question can be answered more effectively with respect to all three data usage scenarios.

1.1 Definition of "personal data"

In this first step, it is not (yet) crucial to decide which data protection law actually applies. As mentioned in the introduction, several laws are in question, such as the ePrivacy Directive (or the upcoming ePrivacy Regulation), the Directive for the police and criminal justice sector or the GDPR. The reason for this is that all laws apply if the processed data is "personal". With regard to this term, all laws basically apply the same definition as under Art. 4 sect. 1 GDPR.[52] Recital 26 GDPR, as well as the Art. 29 Data Protection Working Party give further guidance on the interpretation of the term, which has meanwhile been also taken over by the ECJ.[53]

1.1.1 Step one: Information relating to an identified or identifiable person (the "content", "purpose" and "result" element)

Under Art. 4 sect. 1 GDPR the following is stated:

> "'personal data' means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person".

The following three elements proposed by the Art. 29 Data Protection Working Party can be considered to answer the question of whether information *relates* to an identified or identifiable

[51] Schantz in Wolff/Brink BeckOK DatenschutzR DS-GVO Art. 1 cip. 6; BVerfGE 65, 1 (43) – Volkszählung = NJW 1984, 419 (422)".
[52] See Art. 2 ePrivacy Directive; Art. 4 Sec. 1 (a), recital (5) Proposal for the ePrivacy Regulation; for the relation between the term "personal data" and "electronic communication data" see also ZD-Aktuell 2017, 05452 and ZD 2017, 251; Art. 3 (1) Directive EU 2016/680 ('Directive for the police and criminal justice sector').
[53] See ECJ C-434/16 (Nowak), cip. 35 et seq.

individual: the "content" element, the "purpose" element and the "result" element.[54]

The "content" element refers to the situation where information is *about* an individual. This does not mean that it must be easy for the controller to identify the individual. Rather, the controller may only have some details at the categorical level (e.g. a common name, plus nationality and age) and therefore, has to do quite some research to find out who is actually behind it.[55] The "purpose" element addresses a situation "when the data are used or are likely to be used (...) with the purpose to evaluate, treat in a certain way or influence the status or behaviour of an individual."[56] Even if the "content" and "purpose" element is absent, the data can still relate to an individual because of the "result" element. This element refers to situations where the use of the data has or is likely to have a de facto impact on the individual. Concerning this point, the Working Party outlines that this does not have to be a major impact. Instead, they argue that it is sufficient if the individual is treated differently to other data subjects because of the use of the data.[57]

Given these criteria, information relates to an *identified* natural person if the identifier used is sufficient to uniquely identify that person.[58] For example, if a person's name is unique, information relating to this name refers to an individual who is definitely identified.[59] In contrast, if the identifier *per se* does not suffice to uniquely identify the individual (i.e. single them out), because the name is very common for example, the information relating to that name does not relate to an identified individual. In such a case, it depends on further factors to answer the question of whether that information relates at least to an *identifiable* individual.[60]

1.1.2 Step two: The "means reasonably likely to be used"

With regard to the question of the conditions under which the data must be considered to relate to an *identifiable* individual, recital 26 of the General Data Protection Regulation contains further guidance as stated in sent. 3 and 4 of recital 26:

> "To determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly. To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for

[54] See Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007, 01248/07/EN, WP 136, URL: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf, http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp136_en.pdf, pp. 10 and sequ.
[55] See Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007, 01248/07/EN, WP 136, URL: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf, http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp136_en.pdf, p. 13.
[56] See Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007, 01248/07/EN, WP 136, URL: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf, http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp136_en.pdf, p. 10.
[57] See Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007, 01248/07/EN, WP 136, URL: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf, http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp136_en.pdf p. 11.
[58] The concept of "singling out" an individual "from a group of data subjects" is only problematic insofar as it neglects a negative impact caused by the processing of data, which is presented in *v. Grafenstein, M., The Principle of Purpose Limitation: The Risk-Based Approach, Legal Principles and Private Standards as Elements for Regulating Innovation, 1st ed. (2018)*, Mohr Siebeck, pp.365.
[59] See Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007, 01248/07/EN, WP 136, p. 10. URL: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf
[60] See Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007, 01248/07/EN, WP 136, p. 13. URL: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf

> identification, taking into consideration the available technology at the time of the processing and technological developments."

Referring to the "means reasonably likely to be used", the law applies with a rather flexible approach, which makes it possible to react accordingly to the particularities of a specific case. In this regard, the Art. 29 Data Protection Working Party emphasises that costs and time are not the only relevant factors. Also "the advantage expected by the controller" and "the interests at stake for the data subjects" can and should be taken into account, as well as the purpose of the controller to identify an individual.[61] The reason for the latter factor is that it would be elsewise contradictory if the data protection law were not applicable, even if an "attacker" wants to identify the individual.[62] Only if the means are illegal, the ECJ decided that the means are not "likely reasonably to be used".[63] The rationale for this seems to be that there is no need for additional protection by the data protection law if such an (state) action is already forbidden by another law.

1.1.3 Step three: Anonymisation and pseudonymisation (data minimisation)

Only if data "does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable", the GDPR does not apply; this can be in particular the case when anonymous data is processed for statistical or research purposes.[64] Indeed, to successfully anonymise personal data is a difficult task, at least according to the proposed criteria by the Art. 29 Data Protection Working Group. The Group assesses successful anonymisation according to the following three questions: 1) Is it still possible to single out an individual? 2) Is it still possible to relate data to each other that relate to the same individual? And 3), is it still possible to infer conclusions about an individual from data? If only one of these questions is answered positively, the data is, according to the Working Group's opinion, still personal (i.e. not anonymous). In literature, these three questions have been criticized for their "zero risk" approach.[65] Indeed, in practice, these questions lead to the result that data is considered personal as long as it relates to *an* individual, even if one does not know who this individual actually is. Against this background it becomes clearer what the notion really means, that data is personal as long as one can single out *an* individual – it just does not matter who this is.[66]

Given such a broad scope of application, "pseudonymisation" of personal data can play in practice a more important role than anonymisation of personal data. According to Art. 4 sect. 5 GDPR, "'pseudonymisation' means the processing of personal data in such a way that the personal data can no longer be assigned to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person". Against the aforementioned considerations on personal data and anonymised data, pseudonymised data means

---

[61] See Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007, 01248/07/EN, WP 136, URL: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf, http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp136_en.pdf p. 15.
[62] See Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007, 01248/07/EN, WP 136, URL: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf, http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp136_en.pdf p. 16.
[63] See above in section "II. 1.3.1 Captured IP and MAC addresses (as unique identifiers)", referring to the Judgement of 10/19/2016, Breyer vs. Germany, C-582/14 = ZD 2017, 24 (with annotations by *Klar/Kühling*), pp. 24-25.
[64] See sent. 5 and 6 of recital 26 GDPR.
[65] See, in particular, *El Emam, Khaled, and Álvarez, Cecilia:* A critical appraisal of the Article 29 Working Party Opinion 05/2014 on data anonymization techniques, in: International Data Privacy Law 5 (1) (2015), pp. 73–87.
[66] However, see the more nuanced approach referring to the concerned fundamental right to further determine when an individual is "identified" at v. Grafenstein, Art. 2 DS-GVO, in: Schlender/Stentzel/Veil Kommentar zur DS-GVO, 2nd ed., 2020

that the data is still about an individual but one does not know to whom exactly. This seems to be an important aspect because recital 28 sent. 1 GDPR states that pseudonymisation of personal data "can reduce the risks to the data subjects concerned and help controllers and processors to meet their data-protection obligations."

On this basis, the following sections examine whether the collected data has to be considered as personal data and/or pseudonymised data, thus, whether data protection laws apply, or as anonymised data (all measures used to minimise identification risks are in the following called "data minimisation", see Art. 5 sect. 1 lit. c GDPR). In doing so, the examination seeks to unfold step-by-step the complexity of the question which quality the data fulfils.

1.2 CCTV camera data

Pursuant to Art. 4 sect. 1 GDPR, the collected data via the CCTV cameras is regarded as personal data if, and as long as, the video footage that is being recorded shows an individual (i.e. which can be singled out from a group of further persons in the video).

1.2.1 Recorded data subjects (in particular, faces)

This is undoubtedly the case when the analysis of the video footage results in the identification and/or recognition of individual faces. This data must be considered as data even relating to an *identified* individual.[67] If the records do not show the faces of data subjects, the visual data that had been recorded must be considered as personal (data), as long as somebody else can recognise an individual. In this case, the recorded data relates to an *identifiable* individual. This might be the case because information is gathered through the records that are also combined with further information.[68] For example, it may be enough for family members to recognise the individual based on how he or she behaves. So far, the recorded data is personal because of its "content" element.

1.2.2 Recorded license plates

In the case of the recognition of license plates, this data has also to be considered as personal data, as it refers to an *identifiable* individual. The reason for this is that license plates are registered in the national license plate registers, which also contain the information about the owner of the corresponding vehicle.[69] If the vehicle owner recorded by the CCTV cameras is a natural person and drives this car, that person is directly identifiable due to the number of the license plate. Even if the car owner does not drive the car, the driver is in the most cases indirectly identifiable at the time of recording. The reason for this is that the owner of the car can usually determine who drives his or her car.[70] If not a natural person but a legal entity is the owner of the car, for example, a company, that entity can also determine which natural person drives the car (often by means of a logbook). Therefore, the recorded data can be related to an identifiable individual. These means (i.e., gathering the information from license plate registers and, if required, from vehicle owners) are "reasonably likely to be used" because they require neither significant costs nor time. On the contrary, the license plate registers exist to make such an

[67] See *Klabunde*, in: Ehmann/Selmayr Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 4 cip. 12-13; *Schild* in Wolff/Brink BeckOK DatenschutzR, Art. 4 cip. 14-18; *Dammann*, in: Simitis Bundesdatenschutzgesetz, § 3 cip. 4.
[68] *Klar/Kühling*, in: Kühling/Buchner, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 4 cip. 20-23; *Klabunde* in: Ehmann/Selmayr, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 4, cip. 12-13.
[69] See Klar/Kühling in Kühling/Buchner, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. Nr. 1 cip. 19.
[70] Except for the case that, for example, the car was stolen and driven by an – in this respect – anonymous person.

identification of drivers as feasible as possible.[71]

1.2.3 Obfuscation by the PrivacyProtector (pseudonymisation)

In the technological research project, the provider of the CCTV cameras thinks to "anonymise" the personal data by using the PrivacyProtector, as previously shown.[72] In any case, this could only be the case if the anonymisation process takes place immediately after the recording and when the raw material is irreversibly deleted.[73] However, in the present case, the data may be considered only as pseudonymised, in fact, even if the faces or license plates are obfuscated so that nobody can recognise them. The reason for this is that as long as it is possible to single out an individual in the video footage, irrespective of whether this individual can immediately be identified or not, this individual is still identifiable – for example, because a witness may claim that a recorded (but obfuscated) action of theft has been committed by an individual who he or she has seen on-site. However, before we come to these considerations, we will first develop the technological part of the complexity of the "personal-anonymous data" dichotomy.

1.3 Wi-Fi data

In terms of technology, the situation with regard to the collected data through the Wi-Fi system becomes more complex. In this regard, there are at least two types of collected data that function as an identifier, which means that all other collected data, i.e. time(s) and location(s) of devices, must be considered as personal.[74] The first type of data is the MAC address of a device. As noted above, MAC addresses are used to uniquely identify local network interfaces (e.g. a wireless card). The corresponding 48-bit MAC address under the registration authority of IEEE is assigned to the hardware manufacturers of the respective devices.[75] MAC addresses are used to identify devices within a network. If no MAC address randomisation is applied, the respective device will always identify itself with the same identifier (which is the MAC address) against their network counterpart, e.g. a wireless access point. When an individual accesses the internet with their device, the wireless access point provider assigns a public IP address to a MAC address.[76] This leads us to the second type of identifier (and personal data): IP addresses.

1.3.1 Captured IP and MAC addresses (as unique identifiers)

Both IP and MAC addresses are generally considered as personal data because the link between these addresses and the individual using those addresses is usually stored in a particular place, or at least it is possible to establish the link with a reasonable effort. Both IP and MAC addresses can therefore serve as unique identifiers. The next paragraphs will illustrate, in more detail, why.

Regarding IP addresses, those addresses can relate to an identifiable individual because the internet access provider usually authenticates the user of its network, for example, if an individual logs into the public Wi-Fi network (e.g. giving his or her name, postal address and/or banking details). In this case, the Wi-Fi access provider can refer the IP address to that *identified* individual. However, also from the perspective of third parties, these types of data can refer to an at least *identifiable* person. The reason for

[71] See Klar/Kühling, in: Kühling/Buchner, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 4 Nr. 1, cip. 30; SVR 2014, 201 (204).
[72] See above in section "I. 3.1.3 CCTV camera data collected ('anonymised')".
[73] Cf. BVerfG, Kennzeichenerfassung, NJW 2008, 1505; Gola/Klug/Körffer in Gola/Schomerus, BDSG § 3 cip. 40 ff.
[74] Regarding further unique identifiers, see Eisele et.al. Forum Privatheit, White Paper Privatheit in öffentlichen WLANs, 2017, pp. 62 seq.
[75] Cf. http://standards.ieee.org/develop/regauth/oui/index.html.
[76] See Eisele et.al. Forum Privatheit, White Paper Privatheit in öffentliche WLANs, 2017, pp. 19-20.

this is that such a third party can always contact the entity having access to the authentication data (e.g. the internet access provider) and ask that entity for the authentication information. For example, a website provider who only knows the IP address of a person that is using the website (and the associated usage data such as time and duration) can ask the internet access provider, where the individual left their authentication data. Ultimately, it is actually sufficient that just one service provider is able to authenticate the individual using an IP address to provide the connection. For example, if an internet user made a purchase at a webshop and left their IP address plus their name and billing address, another provider who has only the IP address has just to find this webshop to make the connection. In these cases, it therefore depends on how likely reasonable it is for third parties to find such an authenticating entity. As mentioned previously, this raises many subsequent questions. So far, the ECJ has, at least, made clear that an illegal access to this authentication information cannot be considered as a "means reasonably likely to be used".[77]

Likewise, MAC addresses are personal data whenever it is possible to associate them with an individual, for example, when an internet access provider (assigning the IP address to the MAC address) authenticates the user directly (e.g. by his or her name, postal address and/or banking details). In this case, the provider knows the connection between the user and his or her MAC address. However, this can also happen here indirectly. The moment an (online) service provider learns a device's MAC address, that provider only needs the appropriate additional information to make the connection (e.g. by contacting another entity that has the connection). In these cases of indirect authentication, it depends again on how reasonably likely it is for third parties to find the authenticating entity. This is possible, at least, as long as this access is not illegal.[78]

1.3.2 Truncating and hashing of addresses (pseudonymising unique identifiers)

To avoid the situation where a third party can connect an IP and/or MAC address to an individual via another "authentication entity", it is possible to obfuscate these addresses. Two techniques commonly used in this regard are truncating and hashing. Truncation means that certain parts of an IP or MAC address are deleted. Regarding IP addresses, truncation plays an essential role in preventing any IP address processing from falling within the scope of data protection law(s). Meanwhile, there is a commonly accepted rule that IP addresses are considered to be anonymised when the last eight bits of the addresses are deleted ("considered" because this shows that the question of whether an individual is identifiable via such a truncated IP address or not, is also a normative question depending on the view of certain people, like judges or people working for a data protection authority).[79] To understand this rule and so the more general problem, it is necessary to delve a little bit deeper into the technical details.

IP addresses consist of 32 bits (in its binary representation) with respect to IPv4. An example for such an IP address are the following two sequences:[80]

Decimal presentation	Binary presentation
171.15.245.1	10101011 00001111 11110101 00000001

[77] See the Judgement of 10/19/2016, Breyer vs. Germany, C-582/14 = ZD 2017, 24 (with annotations by Klar/Kühling), pp. 24-25.
[78] See the Judgement of 10/19/2016, Breyer vs. Germany, C-582/14 = ZD 2017, 24 (with annotations by Klar/Kühling), pp. 25-26.
[79] See Kühn, U., "Geolokalisierung mit anonymisierten IP-Adressen" in Datenschutz und Datensicherheit (DuD), 2009 (33), p. 474-751, available under: https://www.datenschutz-hamburg.de/uploads/media/Geolokalisierung_mit_anonymisierten_IP-Adressen__DuD-Beitrag_.pdf.
[80] See the following site: https://www.datenschutzbeauftragter-info.de/ip-adressen-funktion-aufbau-tracking/.

After truncation of the last eight bits	
171.15.245.~~1~~	10101011 00001111 11110101 ~~00000001~~

The truncation of the last eight bits results from a compromise between two conflicting goals: On the one hand, the truncation should make the localisation of an IP address statistically impossible, so that an individual behind that address can no longer be tracked (for example, for advertising purposes). The reason for this is that the localisation of the IP address can be so precise that the individual using that IP address is considered as being identifiable (because in the most extreme case, one could go to the precise location and look who's living or working there). On the other hand, the truncation should not go beyond this goal and make all kinds of geolocation impossible. If only the last eight bits are deleted, it is still possible to find out in which region the address is used. This is important information for adapting websites, for example, to the regional language or applicable laws.[81]

In contrast, MAC addresses are frequently (additionally) hashed. For example, in the current technological research project, Cisco, the provider of the Wi-Fi system, uses a one-way hash function in combination with further techniques, such as truncation. This might be considered as an "anonymisation" of the collected MAC addresses. The reason for this is that the hashing of MAC addresses and the truncation of the hash-value can reduce the connectivity between these addresses and people. To understand this consideration, it is necessary to take, again, a closer look at how these techniques specifically work:

First of all, it is important to note that hashing cannot be used if an *internet* access provider wants to connect a device to the internet because the provider needs to know the MAC address.[82] However, hashing can be used for *online* services provided over the Internet connection, as these services only need their provider to re-identify the device (but not the MAC address). Thus, a hash allows a service provider to re-identify a device without permanently storing the MAC address. This process can be described as follows:

1. A device is detected by a router by initiating communication, which includes the MAC address.
2. The router records the MAC address but immediately hashes the address by using a cryptographic one-way hash-function and immediately deletes the MAC address afterwards.
3. The one-way hash-function creates a specific hash-value. Only this hash-value is stored.

Such a one-way hash means that the original MAC address can in principle, not be derived from the hash-value itself. However, since a certain MAC address will always result in the same hash-value, the hash-value, as already mentioned, allows the device to be re-identified. This leads to two subsequent re-identification risks. Firstly, re-identification is feasible by simply assigning "known" MAC addresses to the hashes. Thus, an entity that has access to the hashes and to a number of identifying MAC addresses can map the hashes to the MAC addresses by re-calculating the hash values of all known MAC addresses and associating the latter with an individual (such as previously described). Furthermore, since the MAC

[81] See Kühn, U., "Geolokalisierung mit anonymisierten IP-Adressen" in Datenschutz und Datensicherheit (DuD), 2009 (33), p. 474-751, available under: https://www.datenschutz-hamburg.de/uploads/media/Geolokalisierung_mit_anonymisierten_IP-Adressen__DuD-Beitrag_.pdf.

[82] For the same reason, IP addresses cannot be hashed because IP addresses are necessary for sending and receiving data packages via the internet, which is outlined in the above in section "I. 2.2.1 The system overall".

address only operates in a (relatively small) 48-bit space,[83] a re-identification could also be conducted by calculating the hashes for all possible MAC addresses with reasonable effort. Both risks can only be reduced by adding a so-called "secret" (commonly referred to as "SALT") to each MAC address before hashing. Such a SALT is a random number. As long as an attacker does not know this SALT, he or she cannot re-calculate the hash-value to a specific MAC address.

As a result, hashing of MAC addresses cannot be considered as a complete anonymisation of the address because the hash value is equally an identifier (which refers to one device and, thus, likely to one carrier of the device). However, the hashing process can be considered as a means of pseudonymisation. If the Wi-Fi access provider ensures that third parties (e.g. online service providers) only get the data related to a hash-value, but not the SALT or even the MAC address itself, such as third parties cannot find out, at least not via this hash value, who the individual is which uses the device.[84]

1.3.3 Movement patterns and randomization of (MAC) addresses

Even if MAC addresses are effectively pseudonymised (including a separately stored SALT), there is still the risk that the user of a device can simply be (re-)identified by other information. This is the case in particular for movement patterns that relate to an identifier, regardless of whether this identifier *per se* does not allow the person to be re-identified. The reason for this is that movement patterns are generated by the collection of location data over a certain period and the more accurate this movement pattern becomes, the more likely it becomes that this pattern reveals information that discards an individual. The following line of thought might help understand this risk: The information that a device moves every morning from location A to location B and back in the evening makes it very likely that this pattern refers to a person on their daily commute to work. Combined with further information, such as (if one just goes there to see) who lives in that single-family home at location A, it is relatively easily possible to find out who this person is.

Hashing MAC addresses does not change this conclusion. As explained before, even if MAC addresses are hashed, it is still possible to single out a personal device and, thus, collect the locations of that device. This data collection creates not only a movement pattern of the personal device but also of the individual carrying the device.[85] In principle, to reduce this risk, i.e. generating such a movement pattern by tracking the device, it is necessary to renew the hash after a certain period. The moment the hash is renewed (and no link to the older hash is left), in theory, one cannot connect anymore the generated movement patterns to each other. The same result can be achieved by MAC address randomisation, as mentioned in the introduction to this section. In this case, the manufacturer has configured the device, more specifically, the local network of the device, so that the MAC address automatically changes after a certain time. Thus, the movement pattern only refers to the period in which the specific MAC address or hash was used unless other means of re-identification are applied.[86]

However, also this approach bears two further pitfalls: First, also the movement pattern per se can serve as a unique identifier. The following example illustrates this idea: Suppose that only one device moves from A to B and back in the morning every morning when a particular MAC address or hash is used.

---

[83] See the following site, as it is in fact only 38,22 bit: https://threatpost.com/research-finds-mac-address-hashing-not-a-fix-for-privacy-problems/104893/.

[84] Cf.Stentzel/Jergel, Art. 4 Nr. 5, cip. 6-7, in: Schlender/Stentzel/Veil Kommentar zur DS-GVO, 2nd ed., 2020 with further requirements regarding the technical and organisation procedures for an effective pseudonymisation.

[85] Pre-supposed that it is always the same individual carrying the device.

[86] For example, using the fingerprint of the technical information that is sent as part of the probe requests from a device's Wi-Fi client to the Wi-Fi access point, see the following site: http://papers.mathyvanhoef.com/asiaccs2016.pdf. Furthermore, the possible countermeasures are listed in the same article.

After a period of time, the MAC address or hash is renewed, but the new address or hash will show the same movement pattern. In this case, it is likely that the movement pattern *as such* still refers to the same device (or to the individual carrying that device). Thus, the pattern itself serves as an identifier which makes it possible to connect those patterns with each other. The second pitfall is that even the sole information that a particular device was in one place at a given time might be enough to identify the person carrying the device. This could simply be the case because another person saw the person on the spot. Thus, renewing hashes or randomising MAC addresses does not mean that there is no risk at all, but only that the re-identification risk is reduced to a certain level. As a rule of thumb, the shorter the time span in which the MAC addresses or hashes are renewed, the lower the risk of re-identification. Comparable to the truncation of IP addresses, therefore, the balance must be found between reducing this risk of re-identification and the usefulness of the collected data for the purposes of the data processing, such as for urban traffic management.

1.4 Parking lot sensor data

The following examination of the collected data by the parking lot sensors shows an illustrative example of how only one piece of information about a place can relate, or in other words, lead to an individual. In doing so, it is very useful to take also the other data usage scenarios into account. In fact, further processing purposes apart from "urban traffic management" are not part of this legal assessment. However, concerning the question whether or not the collected data should be considered personal, it is only relevant that there is at least one situation in which the collected data can relate to an individual. In this regard, the other two scenarios "parking lot assistant" and "law enforcement" make this hypothetical approach quite clear.

1.4.1 Collected data about the parking spots and objects standing on them

At a first glance, the legal assessment does not appear to be complicated. As previously stated, the sensors only collect the following data:

- Size of a parking lot in a predetermined area;
- GPS-position of a parking lot;
- and, whether there is an object in this parking spot (i.e. whether a vehicle parks there or not).

At a first glance, this data does not relate to an *identified* person (such as the collected data via the CCTV cameras before their obfuscation). The data does also not relate to a unique identifier that refers to an identifiable individual (however, the MAC addresses of the devices that are collected via the Wi-Fi system relate to identifiable data subjects). But in a second view, even the collected data of the parking lot sensors can relate to at least an identifiable individual. The two scenarios "parking lot assistant" and "law enforcement" illustrate where, how and why they relate to an individual.

1.4.2 Data collected during the use of the parking lot assistant

As illustrated previously, in the second data usage scenario of the current DPIA, a user of a personal device connects to the public Wi-Fi system to use the parking lot assistant provided by Cleverciti. The user has to create a personal account to be able to choose and find an available parking space and pay for it if it is necessary. The moment the user connects to the Wi-Fi system, Cleverciti will be able to

associate all data that is collected during the use of that service to this specific user. This data contains information about the location, time and on how the user interacts with this service.[87]

There are several reasons why this data relates to the user. The first reason refers to the "content" element: First, if the user has to pay for a parking space, the payment details provided by the user contains authentication data. Otherwise, payment would not be possible.[88] Second, the email address registered under the personal account could also authenticate the user.[89] Even if the user has not paid nor registered his or her email address (which authenticates the individual), it would be possible to identify the user. The reason for this, as already shown, is that the knowledge of MAC and/or IP allows Cleverciti to combine these identifiers with other information, such as authentication data that the user may have given to another online service provider. Again, in this regard, the question is whether such additional information is "reasonably likely to be used". Taking "all objective factors" into account, it may indeed be too costly and/or too time consuming for Cleverciti to identify the user by combining his or her MAC and/or IP address with additional information (that is possibly kept by another entity).

In these cases however, the data could relate at least to the users due to the "purpose" and "result" elements: for example, if Cleverciti wishes to treat the user in a certain way, i.e. to guide them to a free parking slot, then the mere "purpose" of Cleverciti results in the application of data protection law. Beside the "purpose" element, also the "result" element can still lead to the application of the law. Such a "result" element could be at stake if the data of the users are used in a "discriminatory" way. For example, if Cleverciti requires users to pay for a parking spot reservation, non-paying users may "suffer" from the impact of not getting the desired parking lot. If Cleverciti uses the general interest of such non-paying users in certain parking spots to signal to users who are willing to pay the "urgency" to reserve the parking spot, this information is used against the non-paying users. Since the "result" element does not require a major impact on an individual but only that he or she is treated differently compared to other users, in the case of a required fee, the "result" element applied.[90]

1.4.3 Data collected irrespective of the use of the parking lot assistant but combined with further information (in particular, "on-site")

Given the third data usage scenario, even all the collected data by the parking lot sensors can be related to (at least) identifiable data subjects. As previously described, in the first alternative of the third data usage scenario a public order agency uses the dashboard that is offered by Cleverciti to mark certain areas as restricted (according to information from the local traffic street regulation). On this basis, the agency can observe via the cockpit, which is an additional service offered by Cleverciti, where vehicles stop in prohibited zones. Receiving this information at the moment when xyz is happening, enables the public order agency to immediately send its security staff to the site. At the actual location of the offence, the security staff writes down the license plate number to refer the previously "anonymous" data to an individual, i.e. the owner of the vehicle.

When the dashboard and/or cockpit functions are in use enabling the public order agency to fulfill its traffic law enforcement tasks, the "purpose" of identifying the owners of those cars parked in prohibited areas leads to the application of data protection law. There are only two reasons that could disagree with

[87] See above in section "I. 3.2.2 Parking lot assistant (based on user's consent)".
[88] Although blockchain-based currencies, such as bitcoin, could be an option for anonymous payment transfers if further conditions are met, which can be found in https://www.privacy-handbuch.de/handbuch_26_bitcoin.htm
[89] Cf. Brink, S., Eckhardt, J.: "Wann ist ein Datum ein personenbezogenes Datum? - Anwendungsbereich des Datenschutzrecht" in ZD 2015, 205s; Ernst in; Paal/Pauly, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 4 Sec. 1 lit. 1 cip. 14.
[90] Cf. Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007, 01248/07/EN, WP 136, URL: http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp136_en.pdf, p. 11.

such a conclusion: First, if the collected data is displayed on the cockpit screen directly after it has been collected and then immediately and irreversibly deleted, one can doubt if there are data protection risks at all. The reason for this doubt is that some constitutional courts essentially consider a need for data protection if the data is *digitally* stored. The German Constitutional Court considers, for example, that the digital storage makes it possible to "retrieve the data, in a matter of seconds, at any time from any place" and to "combine it with further data collections creating partly or comprehensive personality profiles" as well as "various possibilities to use it".[91] Likewise, the European Court of Human Rights sees the particular danger in the systematic and permanent storage of data[92] and, consequently, refuses protection if a surveillance camera records data for screening purposes only, but does not store it.[93] However, the GDPR makes clear that already the collection, transmission, and screening of personal data are already within its scope of protection, regardless of whether the data is stored or not.[94] Therefore, when the data is recorded, transmitted and screened to identify a person, the GDPR applies.

Another reason against the application of data protection law could be that the relation of on-site parking space data, i.e. the situation in which a person breaks the traffic road law, could be regarded as illegal and therefore, a means that is not "reasonably likely to be used".[95] The question is, hence, whether the public order agency is allowed to use the collected data by the parking lot sensors for this purpose. According to Art. 46 of the German Act on Regulatory Offences, Art. 98 lit. c of the German Code of Criminal Procedure could apply to such a situation and, therefore, serve as a basis for the data processing.[96] However, pursuant to Art. 47 of the German Act on Regulatory Offences, the "prosecution of regulatory offences shall be within the duty-bound discretion of the prosecuting authority." In our case, this provision means that if the public order agency does not use its margin of discretion the prosecution is affected by a procedural deficiency. This could be the case if the whole process of traffic road law prosecution is automated.[97] In our case, however, it was assumed that the public order agency sends out a person responsible for the law enforcement and who can also use their discretion on the ground. Therefore, this process is not affected by a procedural deficiency, and thus, is not an unlawful means which would lead to an exclusion of the data protection law.[98]

However, the situation might be different with respect to the second alternative of the third data usage

---

[91] See BVerfG, Urt. v. 15.12.1983 – 1BvR 209, 269, 362, 420, 440, 484/83 (Volkzählungsurteil), cip. 153, and BVerfG, Urt. v. 4.4.2006 – 1BvR 518/02 (Rasterfahndung), cip. 65: "... technisch gesehen unbegrenzt speicherbar und jederzeit ohne Rücksicht auf Entfernungen in Sekundenschnelle abrufbar sind", dabei „mit anderen Datensammlungen zu einem teilweise oder weitgehend vollständigen Persönlichkeitsbild zusammengefügt werden" und „vielfältige Nutzungs- und Verknüpfungsmöglichkeiten entstehen".

[92] See ECHR, P.G. and J.H. vs The United Kingdom vom 25.9.2001 (application no. 44787/98), cip. 57 referring to ECHR, Amann vs Switzerland [GC], no. 27798/95, §§ 65 bis 67, ECHR 2000-II, and Rotaru vs Romania [GC], no. 28341/95, §§ 43-44, ECHR 2000-V.

[93] See ECtHR, Case of P.G. and J.H. vs. The United Kingdom from 25 September 2001 (application no. 44787/98), cip. 57: "Since there are occasions when people knowingly or intentionally involve themselves in activities that are or may be recorded or reported in a public manner, a person's reasonable expectation as to privacy may be a significant, although not necessarily conclusive, factor. A person who walks down the street will, inevitably, be visible to any member of the public who is also present. Monitoring by technological means of the same public scene (for example, a security guard viewing through closed-circuit television) is of similar character. Private-life considerations may arise, however, once any systematic or permanent record comes into existence of such material from the public domain. It is for this reason that files gathered by security services on a particular individual fall within the scope of Article 8, even where the information has not been gathered by an intrusive or covert method (...)".

[94] See *Herbst* in Kühling/Buchner, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 4, cip. 21; Ernst in Paal/Pauly Art. 4 cip. 23 and 30.

[95] See above in section "II. 1.3.1 Captured IP and MAC addresses (as unique identifiers)", referring to the Judgement of 10/19/2016, Breyer vs. Germany, C-582/14 = ZD 2017, 24 (with annotations by Klar/Kühling), pp. 24-25.

[96] See *Graf* in Bücherl BeckOK OWiG, § 46 cip. 1, Lampe in Karlsruher Kommentar zum OWiG, § 46 cip. 4, Greven in Karlsruher Kommentar zur StPO, § 98 c, cip. 1-3.

[97] Cf. *Graf* in Bücherl BeckOK OwiG, § 47, cip. 7-13; Bohnert/Krenberger/Krumm, OWiG § 47, cip. 5-6.

[98] See the judgement of 10/19/2016, Breyer vs. Germany, C-582/14 = ZD 2017, 24 (with annotations by Klar/Kühling) ECJ C-582/14 (Breyer), cip. 46.

scenario. In this second alternative, in a first step, a public order agency gains access to the collected data by the CCTV cameras, which is immediately "anonymised", to monitor whether "anonymous" persons are crossing a street to which they are not permitted under current traffic law. As soon as the public order agency discovers such a "delinquent" (yet "anonymous") person, the agency, as a second step, uses its access to the collected data via the Wi-Fi system. This enables them to find out which mobile device that is connected to the Wi-Fi system exactly matches with the delinquent person. This can certainly happen if the "delinquent" person is logged into the Wi-Fi network and the user has previously authenticated him or herself. In contrast, if the "delinquent" person has not authenticated but can only be indirectly identified (e.g. via the MAC address), it depends on whether such additional identifying information is "reasonably likely to be used". If this is the case, the public order agency can finally identify that person who crossed the street illegally. To make the whole process more cost efficient, it is completely automated.[99] Similar to the first alternative of this scenario, the public order agency wants to identify the individual. However, unlike the first alternative, in this second alternative the process is completely automated and can therefore lead to a procedural deficiency if the margin of discretion is not used as required. In this case, the combination of the two different datasets may be an illegal means, which means that data protection law is not applicable.[100]

1.5 Interims conclusion

In light of the above considerations, the current DPIA concludes that all the collected data in the framework of the Smart City research project must be considered as personal data for a variety of reasons. The following list provides a brief summary of the results:

- *The data collected by the CCTV cameras* must be considered as personal data, firstly, as long as faces of data subjects (relating to an identified individual) and/or license plates are recorded (relating to an identifiable individual); secondly, these data can only be considered pseudonymised if their obfuscation makes it impossible for a controller to directly identify the individual (e.g. by a family member) or the license plate, without additional information. In contrast, the data can only be considered anonymised if their obfuscation makes it impossible to say that there is *an* individual recorded (in this last regard, see the example given at the end of this interims conclusion), for example, because there may also be an indefinite number of recorded and obfuscated individuals.
- *The data collected by the Wi-Fi system in particular the MAC addresses* must be considered as personal data (relating to an identifiable individual); if the MAC addresses were hashed and a SALT were to be added in a second step, the data could be considered pseudonymised data (but not anonymised data). Especially regarding movement patterns, there is a generally high risk that the location data could be related to an individual by adding other information (e.g. a person recognizing the movement pattern of his or her partner, moving from home to work and back). This risk can be reduced by, for example, renewing the hashes or by randomising MAC addresses, but it can hardly be excluded. Comparable to the truncation of IP addresses, the determination of the interval after a hash or MAC address renewal must balance the corresponding re-identification risk for the data subject and the research interest in the analysis of these movement patterns in the area of urban traffic management.
- The data collected by the *parking lot sensors* must also be considered personal. The reasons for this are twofold:
    - In the first (simple) case, the collected data in relation to the parking lot assistant can

[99] See above in section "I., 3.2.3 Law enforcement".
[100] See ECJ decision of 10/19/2016, Breyer vs. Germany, C-582/14 = ZD 2017, 24 (with annotations by Klar/Kühling).

refer to data subjects if two conditions are met: The data subjects create a personal account (somewhere) that contains data enabling the provider (or other parties) to (indirectly) identify them; and moreover, these two different types of data are combined. Even if there were no such authentication data anywhere where it could reasonably be found on the internet, the data could nevertheless be personal since the data subjects are likely to be treated differently as a result of the data processing (e.g. a user paying for a parking lot might reserve it, while others will not, but the information about the interest of the non-paying user for parking lots is feeded into the system).

- Second, even all the collected data can relate to data subjects, regardless of the car park assistant, if combined with appropriate additional information. For example, if a public order agency sends a person to the place where the data is recorded and compares it with information on the site (e.g. by writing down the license plate number and matching this number with the stored information about the car-owner in the license plate registry).

The last two examples demonstrate that there is a very high risk that all collected data relate to an individual if it is combined with other appropriate information. This consideration even applies to the collected data by the CCTV cameras, which is immediately obfuscated so that nobody can recognise the individual in the footage (i.e. who this is). Even if it is impossible to directly identify a person, who commits an offence (or even a crime) recorded in such a video, the mere information that this someone has committed such an offence can be used against that person. This gets clearer if one imagines the situation, for example, when a witness cannot say what exactly happened in a particular place and time, but he or she can testify that a certain person was there at the time. Both pieces of information, i.e. about the action (that was recorded by the cameras) and the person (who was seen by the witness) can then be combined later. This seems to be the reason for the "zero risk" approach of the Art. 29 Data Protection Working Party that sees a de-anonymisation risk already when it is possible to single out an individual – even if one does not know in the beginning who this individual is.[101] However, these examples illustrate why, in fact, there is little data that could be considered non-personal or completely anonymised because there is always, at least with respect to the data usage scenarios, a way how actually non-personal data can be referred to an individual. As a consequence of this rather broad approach, this DPIA comes to the conclusion that all data collected by the system must be considered as personal. Of course, the previous analysis has made it also clear that the collected data entails different degrees of pseudonymisation and thus different risks of relating or being related to an individual. There are also different degrees of sensitivity of the recorded information (for example between a normal passer-by situation, a crime committed in public, or when somebody has to cross naked a street because he locked himself out of his apartment), irrespective of how this information is finally used. These different risks must be taken into account when defining the appropriate data protection by design strategy.[102]

2 Necessity and Proportionality

Given that data protection laws apply, the data processing must be necessary to achieve the purpose of the processing, and also has to be proportionate. The first requirement that the controller(s) has (or have) to meet, in this regard, is the specification of its purpose for the data processing. Only in the moment

[101] See Article 29 Data Protection Working Party, Opinion 5/2014 on Anonymisation Techniques, adopted on 10 April 2014, 0829 /14/EN WP 216, pp. 11 and 12.
[102] See Grafenstein, Art. 2, in: Schlender/Stentzel/Veil Kommentar zur DS-GVO, 2nd ed., 2020; Stentzel/Jergel, Art. 4 Nr. 5, cip. 6-7, in: Schlender/Stentzel/Veil Kommentar zur DS-GVO, 2nd ed., 2020

the purpose is specified, it is possible to determine which data protection law applies. Only on this basis, it is possible to examine whether or not the processing is lawful, and how further principles apply, such as the principles of purpose limitation and storage limitation.

2.1 Purpose specification: "Research and statistics for urban traffic management"

Pursuant to Art. 5 sect. 1 lit. b GDPR, the controller is required to specify the purpose of its data collection and later processing activities. This purpose should be, pursuant to recital 22 sent. 6 GDPR, "explicit and legitimate and determined at the time of the collection of the personal data." One of the major questions regarding this requirement is how precisely the controller has to specify the purpose.[103]

2.1.1 Challenge when specifying research and statistic purposes

This is particularly the case with regard to the data usage scenarios as described in this current DPIA. As already explained, the target of evaluation of this DPIA cannot (yet) be determined given the preliminary development phase of the proposed system. Therefore, the current DPIA has chosen and described three scenarios how the data could be collected within the framework of the technological research project and could be used later on. In view of this situation, the question arises as to how exactly the inherent processing purposes of these scenarios can be specified from a factual standpoint. This means, taking into account the first data usage scenario: Is the purpose of "research and statistics in the area of urban traffic management" sufficiently precise?

2.1.2 Discussion on sufficiently precise specification

The Art. 29 Data Protection Working Group has given some guidance on this question as in regard to the currently applicable Data Protection Directive:

> "Personal data can be collected for more than one purpose. In some cases, these purposes, while distinct, are nevertheless related to some degree. In other cases, the purposes may be unrelated. A question that arises here is to what extent the controller should specify each of these distinct purposes separately, and how much additional detail should be provided. (...) For 'related' processing operations, the concept of an overall purpose, under whose umbrella a number of data processing operations take place, can be useful. That said, controllers should avoid identifying only one broad purpose in order to justify various further processing activities which are in fact only remotely related to the actual initial purpose. Ultimately, in order to ensure compliance with Article 6(1)(b), each separate purpose should be specified in enough detail to be able to assess whether collection of personal data for this purpose complies with the law, and to establish what data protection safeguards to apply."[104]

Given these considerations, the purpose of research and statistics for urban traffic management is indeed quite broad. However, there are several reasons for why this purpose may nevertheless be sufficiently precise: First, this purpose serves as an umbrella that does not 'justify various further processing activities which are in fact rather remotely related to the actual initial purpose.' Instead, the purpose of research and statistics for urban traffic management summarises the different processing activities, and consequently, sub-purposes that are necessary to achieve this aim. Second, research and statistics are

[103] See *Schantz* in Wolff/Brink BeckOK DatenschutzR, Art. 5 cip. 15-17; Dammann in ZD 2016, 307.
[104] See the Article 29 Data Protection Working Party, 00569/13/EN WP 203, adopted on 2 April 2013: Opinion on purpose limitation, p 16.

inherently open-ended. Therefore, it is difficult, if not impossible, to limit them to more specific subgoals.[105] Recital 33 of the GDPR argues in favor of the purpose to be sufficiently precise:

> "It is often not possible to fully identify the purpose of personal data processing for scientific research purposes at the time of data collection. Therefore, data subjects should be allowed to give their consent to certain areas of scientific research when in keeping with recognised ethical standards for scientific research. Data subjects should have the opportunity to give their consent only to certain areas of research or parts of research projects to the extent allowed by the intended purpose."

If recital 33 GDPR recognises that data subjects can even *consent* to certain research areas (not fully identifying more precise sub-purposes), then it should be all the more possible to specify such a research purpose in light of Art. 5 sect. 1 lit. b GDPR. The reason for this "*de maiore ad minus*" argument is that a consent provides a legitimate legal basis, pursuant to Art. 6 sect. 1 lit. a GDPR, whereas the requirement to specify the purpose, pursuant to Art. 5 GDPR, is "only" a legal principle.[106] Given that the factual processing activities do not follow other, unrelated (hidden) purposes carried out under the umbrella of "research and statistics for urban traffic management", the current DPIA considers this purpose as sufficiently specific. Last but not least, these research and statistics purposes for urban traffic management are also considered legitimate because research and statistical purposes are mentioned in the law itself.

2.1.3 Supporting transparency measures (Art. 5 sect. 1 lit. a GDPR)

However, to increase transparency for these purposes, this DPIA recommends informing the data subjects as precisely as possible about the execution of the processing operations. Transparency is extremely important to prevent data subjects from suffering from the unspecific threat that they know that certain processing activities are carried out, on the basis of data that is, or was once, relating to them, but do not know what has been, is being or will be done exactly.[107] The current DPIA will later examine in more detail how effective transparency measures can and should be implemented to avoid such a threat.

2.2 Lawfulness of processing

On account of the fact that this DPIA considers the processed data for research and statistical purposes as personal data, the data processing must be based either on the consent of the individual's consent, or on a legal provision. This chapter examines whether the GDPR applies, in this regard, or any other applicable law.

---

[105] Cf. *Nolte* in Schlender/Stentzel/Veil, Kommentar zur DS-GVO, Art. 89, cip. 36.

[106] See the title of this article: "Principles relating to processing of personal data"; see with respect to the "softer" effects of a legal principle, in comparison to the "stricter" effects of a conditional if-then rule; *Eifert*, Regulierungsstrategien, in: Wolfgang Hoffmann-Riem / Eberhard Schmidt- Aßmann / Andreas Voßkuhle (eds.), Grundlagen des Verwaltungsrechts – Band I „Methoden – Maßstäbe – Aufgaben – Organisation", 2nd edition, Mu̇nchen: C.H. Beck, 2012, § 19, cip. 13 to 15; focusing on privacy-related principles, Maxwell, Principles-based regulation of personal data: the case of 'fair processing', pp. 212 to 214, referring to J Black, 'Forms and Paradoxes of Principles Based Regulation', LSE Law, Society and Economy Working Paper 13/2008, SSRN abstract n8 1267722, L Kaplow, 'Rules Versus Standards: An Economic Analysis' (1992) 42 Duke L. J. 557; R Posner, Economic Analysis of Law (8th edn., Aspen/Wolters Kluwer, New York, 2011), p. 747.

[107] See ECJ C-293/12 and C-594/12 (Digital Rights vs. Ireland), cip. 37; and even clearer, BVerfG, 2nd March 2010, 1 BvR 256/08, 1 BvR 263/08, and 1 BvR 586/08 (Data Retention), cip. 241.

2.2.1 Applicable basis laid down by law

The GDPR applies to all types of personal data processing, but there is an exception or another prevailing law. An exception could apply if the data is processed by competent authorities for police enforcement purposes (Art. 2 sect. 2 lit. d GDPR). The third data usage scenario tends in that direction. However, this scenario is not the target of evaluation of this DPIA, but only the processing purposes for research and statistics in an urban traffic management environment. Another exception would apply if the EU institutions were to process the data (Art. 2 sect. 3 GDPR). However, even if it is not yet clear who the data controllers of the intended system are, it is very unlikely that EU institutions are involved. In contrast, in respect to the data collected by the Wi-Fi network, the upcoming ePrivacy Regulation could prevail the GDPR. This could result in a stricter regulation.[108] Only as far as the ePrivacy Regulation does not apply, the processing could be based on the GDPR. In this regard, there are three possibilities that may come into question: First, the processing could be lawful if it is necessary for a task carried out in the public interest (Art. 6 sect. 1 lit. e GDPR). Second, the processing could be based on the individual's consent. And third, finally, the "legitimate interest [- clause]" under Art. 6 sect. 1 lit. f GDPR could apply. After a brief illustration of the aforementioned legal basis, the "legitimate interest ["- clause]" will be the focus of this chapter. The reason for this is that the basic principles discussed with respect to the "legitimate interest ["- clause]" can often be referred to another legal basis.[109]

2.2.1.1 ePrivacy Regulation (Art. 95 GDPR)

Pursuant to Art. 2 sect. 1 of the Proposal for a Regulation of the European Parliament and of the Council concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (in the following "Proposal for an ePrivacy Regulation), this regulation "applies to the processing of electronic communications data carried out in connection with the provision and the use of electronic communications services and to information related to the terminal equipment of end-user." The term "electronic communications data" includes, under Art. 4 sect. 3 lit. a), electronic communications content and electronic communications metadata. Regarding the collected data by the Wi-Fi network, the definition regarding metadata is particularly relevant. Pursuant to Art. 4 sect. 2 lit. e), such metadata is "data processed in an electronic communications network for the purposes of transmitting, distributing or exchanging electronic communications content; including data used to trace and identify the source and destination of a communication, data on the location of the device generated in the context of providing electronic communications services, and the date, time, duration and the type of communication". According to this definition, the data collected by the Wi-Fi network would then fall under the protection of the ePrivacy Regulation.

This result is very important because the data collected by the Wi-Fi network could only be processed for research and statistical purposes if the people had given their consent. This restrictive approach applies to both of the following two situations: First, if a data controller wants to collect the data for research and statistical purposes; second, if the data are used only later for research and statistical purposes but were originally collected for a different purpose. In both situations, a data controller needs the consent of the data subjects concerned. The reason is that there is no legal provision under the

[108] See Engeler, M., Felber, W.: "Entwurf der ePrivacy-VO aus Perspektive der aufsichtsbehördlichen Praxis - Reguliert der Entwurf an der technischen Realität vorbei? in ZD, 2017, pp. 251-257 (253).
[109] See, for example, Art. 2 sect. 3 sent. 2 GDPR stating that the regulation applicable to the processing of personal data by EU institutions shall be adapted to the principles and rules of the GDPR; Art. 6 sect. 3 sent. 3 GDPR, which recommends to specify, on the national legal basis, the general rules of the GDPR; or the responsibility principle and the privacy by design requirement that also apply to this consent.

Proposal of the ePrivacy Regulation that would apply to the data processing. As far as the first situation is concerned, a data controller may in principle only process Wi-Fi data if this is necessary for the technical provision of the communication service, for the maintenance and restoration of the security of the network as well as services or for billing purposes.[110] If the data were originally collected for a purpose other than research and statistics, the data controller also needs the consent of the individual for such change of purpose, as the proposal for the ePrivacy Regulation requires a "purpose identity" rather than a "purpose compatibility".[111] This duty of consent poses a great challenge to data controllers (as in the current Smart City research project). The next chapter will address this challenge in more general terms with regard to the consent under the GDPR.

2.2.1.2 Public interest task (Art. 6 sect. 1 lit. e, and sect. 2 and 3 GDPR in combination with a national legal basis)

Before illustrating the challenges related to the consent, a further legal basis that could be used for research and statistical purposes in an urban traffic management environment should be mentioned. As long as the ePrivacy Regulation does not apply, the processing could be lawful if it is "necessary for the performance of a task carried out in the public interest", pursuant to Art. 6 sect. 1 lit e GDPR. Under Art. 6 sect. 2 GDPR, the EU Member States may lay down such legal basis by specifying the principles established under the GDPR. Such a legal basis may be particularly interesting for research and statistical purposes in an urban traffic management environment because Art. 6 sect. 3 GDPR provides certain legal privileges: First, such a legal basis does not have to specify the purpose of the data processing. Instead, it is sufficient if the legal basis determines the public task. This appears to be a difference to the German situation regarding the right to informational self-determination. This right requires, pursuant to the German Constitutional Court, the legislator to specify the purpose that must be narrower than the public task of a public agency.[112] However, critics emphasise that the legal basis must at least determine the public task so it can be assessed whether processing is necessary or not.[113]

Public service tasks are usually considered to be in the public interest.[114] Comparably, the Art. 29 Data Protection Working Party considers the processing of personal data for smart meterings to be in the public interest, as long as it pursues a more efficient energy supply and energy consumption.[115] This means that the processing of personal data for urban traffic management could also be in the public interest. Furthermore, the national legal basis does not necessarily have to be a parliamentary law, as long as the national law, which can serve as a legal basis for the data processing, is proportionate to the public interest pursued (last sentence of Art. 6 sect. 3 GDPR). At least, recital 41 GDPR makes clear that also other forms of rulemaking come into question, as long as the requirements of the respective national constitutional order are fulfilled and its legal requirements according to the case-law of the ECtHR and the ECJ are so precise that the data subjects concerned are able to foresee its legal

[110] See Art. 6 sect. 1 lit. a and b, sect. 2 lit. b and c of the Proposal for a Regulation on Privacy and Electronic Communications, available under: https://ec.europa.eu/digital-single-market/en/news/proposal-regulation-privacy-and-electronic-communications.
[111] Cf. Art. 6 Sec. 2 lit. c and Sec. 3 ePrivacy Directive
[112] See BVerfG, 20th of April 2016, 1 BvR 966/09 and 1 BvR 1140/09 (Federal Bureau of Investigation Law) = NJW 2016, 1781, cip. 279: "Die jeweilige Eingriffsgrundlage bestimmt Behörde, Zweck und Bedingungen der Datenerhebung und definiert damit die erlaubte Verwendung. Die Zweckbindung der auf ihrer Grundlage gewonnenen Informationen beschränkt sich folglich nicht allein auf eine Bindung an bestimmte, abstrakt definierte Behördenaufgaben, sondern bestimmt sich nach der Reichweite der Erhebungszwecke in der für die jeweilige Datenerhebung maßgeblichen Ermächtigungsgrundlage."
[113] See *Buchner/Petri* in Kühling/Buchner, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 6, cip. 121.
[114] See *Buchner/Petri* in Kühling/Buchner, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 6, cip. 128.
[115] See Article 29 Data Protection Working Party, Opinion 12/2011 on smart metering, 4 April 2011, 00671/11/EN, WP 183, pp. 12 and 13.

consequences. As mentioned previously, Art. 5 sect. 3 sent. 3 GDPR states in this regard that the "legal basis may contain specific provisions to adapt the application of rules of this Regulation, inter alia: the general conditions governing the lawfulness of processing by the controller; the types of data which are subject to the processing; the data subjects concerned; the entities to, and the purposes for which, the personal data may be disclosed; the purpose limitation; storage periods; and processing operations and processing procedures, including measures to ensure lawful and fair processing such as those for other specific processing situations as provided for in Chapter IX."[116] Thus, this DPIA may also provide for guidance under which conditions such a national basis might legitimise the data processing for the research and statistical purposes in an urban traffic management environment.

2.2.1.3 Consent (Art. 6 sect. 1 lit. a GDPR)

Irrespective of the national legal basis that allows the data processing when a task has to be carried out in the public interest, the data controllers could also base the processing on the individual's consent in accordance to Art. 6 sect. 1 lit. a GDPR. However, as mentioned previously in the context of the ePrivacy Regulation, the requirement to base the data processing on the individual's consent would pose a major challenge to the research purposes. The reason for this is that data subjects must give their consent before or at the latest during the collection of the data.[117] In contrast, in the Smart City research project, the data collected by the system is collected before the person can give their consent. With respect to the Wi-Fi data, for example, the earliest moment a person can give their consent is usually in the moment they decide to log-into the network. However, as shown previously, the data collection starts before that moment, i.e. when a Wi-Fi access point receives the probe requests including the MAC address of a person's device and this data is stored.

Likewise, a person is unable to give their consent at the moment CCTV cameras record them or when parking lot sensors detect their car parking in a particular parking lot. Even if this data is immediately obfuscated (such as in the CCTV camera case) or does not record the license plates at all (such as in the parking lot sensor case), this DPIA concludes that this data is nevertheless personal because it *can* be referred, in principle, to an identified individual later on. Since such data is considered personal a legitimate basis is required, and if this legitimate basis is the individual's consent, the individual concerned has to give his or her consent — in the moment of collection. This is the consequence resulting from the hypothetical approach of data protection law, which addresses risks that can lead to harm but do not necessarily lead to harm. Thus, this hypothetical approach becomes already obvious at the first stage of the legal assessment, that is, when defining the term "personal data".[118]

The consent requirement therefore leads to the situation where the collected data may only be processed for research and statistical purposes if and only if the data subjects agree to it beforehand. However, as the consent of the data subjects would be too late and, besides that, the data subjects should have to explicitly opt-in, there also may be too few data subjects giving their consent so that the research and statistical basis might be too small to reach the research purposes in the urban traffic management environment.[119] This is why the question on the application of the ePrivacy Regulation is so crucial. Should the EU legislator neither change the scope of this regulation nor its regulatory approach (i.e.

[116] Regarding measures and requirements in relation to Art. 89 GDPR, as one of the specific provisions for the processing of personal data for research and statistical purposes of Chapter IX, see the points "2.3.3 Legal privileges for research and statistical purposes (Art. 89 GDPR)" and "2.3.3.2 Safeguards required under Art. 89 sect. 1 GDPR".
[117] See *Buchner/Petri* in: Kühling/Buchner, Datenschutz-Grundverordnung, 2. Aufl., 2017, Ar. 6, cip. 18-21.
[118] See in section "II. 1.1 Definition of 'personal data'".
[119] According to the prevailing opinion the consent requirement sees data subjects to opt-in (instead of opt-out). This can be seen, for example in *Frenzel* in; Paal/Pauly, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 6 Sec. 1 lit. a. However, there is a lower degree of probability that people signal their agreement when required to use an opt-in instead of an opt-out procedure.

allowing the processing of Wi-Fi data only for technical, security or billing purposes, or on the basis of the individual's consent), the data controllers of the Smart City research project could only use the collected Wi-Fi data in the moment and if the data subjects agree to it *at all*. In contrast, as far as the GDPR applies, there is at least the chance that the data controllers could also base the data processing on the so-called "legitimate interests"-clause.

2.2.2 "Legitimate interests"-clause (Art. 6 sect. 1 lit. f GDPR)

Whether the processing of the collected data for research and statistical purposes in an urban traffic environment can be based on the "legitimate interests"- clause depends on whether the interests of the data subjects concerned override the research and statistical interests or not. With respect to the processing of Wi-Fi data for mobile location analytics purposes, the Forum Privatheit came to the conclusion that such processing cannot be based on Art. 6 sect. 1 lit. f GDPR. The first main reason for this strict approach was the sensitivity of the data resulting from the individual profiles created to influence the behaviour of the data subjects concerned (i.e. the consumers) and their purchasing decisions. The second reason referred to the risk of a comprehensive surveillance of cities, which is particularly at stake when there are central services offering these analytical services to businesses throughout the city. In such a case, these service providers cannot only track the movement of data subjects within a single shop — restricted to a certain place — but the entire city.[120] Similarly, one could conclude from the decision of "*Mr. González vs. Google Spain*" by the European Court of Justice, that the individual's fundamental rights to private life and data protection under Art. 7 and 8 EuChFR override as a rule not only the economic interest of private companies but also the interests of the public.[121] Of course, this decision referred to a profile of an individual that was created by an internet search engine when users of that search engine typed in the individual's name. The decision, hence, referred to a different case than that of this DPIA. For this reason, one cannot transfer these considerations to the current evaluation without reflecting on them properly. Therefore, beside the commonalities, one must take the particularities of the target of this assessment into account.

2.2.2.1 Legitimate interests of data controller(s)

Pursuant to Art. 6 sect. 1 lit. f GDPR, the processing is only legitimate if it is "necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data". Thus, the first condition requires that the processing is necessary for the purposes of a legitimate interest of the controller or a third party. Since it is unclear who the controller(s) within the intended system is (or are), it is difficult to determine such an interest. However, insofar as private companies are involved — be it as a controller, processor, or manufacturer — these entities can rely, at least, on their fundamental right to conduct a business under Art. 16 ECFR. So far, also public interests come into question and can give more 'weight' to private interests.[122] As mentioned in the above, there are two public interests in this case: Focusing on the statistical purposes, the data processing meets the public interest in an efficient urban traffic management; focusing on research purposes, the data processing meets the interest in increasing a society's technological innovation capacity. Also, recital 157 GDPR explicitly highlights the societal interest in research and statistics. As a consequence, this DPIA considers the research and statistical interests in the data processing, at least, in

[120] See *Eisele* et al., Forum Privatheit, White Paper Privatheit in öffentlichen WLANs, 2017, pp. 56 and 57.
[121] See Judgement of 05/13/2016, González vs. Google Spain, C-131/12, cip. 97.
[122] See Art. 29 Data Protection Working Party, Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC, 9 April 2014, 844/14/EN, WP 217, URL: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp217_en.pdf, p. 35.

an urban traffic management environment as, *per se*, legitimate. Last but not least, given the open-ended nature of research and statistical processes, this DPIA also considers, in principle, the data processing as necessary to satisfy these interests.[123]

2.2.2.2 Interests of data subjects

However, the interests of the involved private companies as well as the interests of the public, must not be overridden by the interests of the data subjects concerned.[124] Also in this regard, and still on the basis of the Data Protection Directive, the Art. 29 Data Protection Working Party has given guidance on how to evaluate the interests of the data subjects concerned. In particular, the Working Party proposes to take the following criteria into account: the context of the data collection; the nature of the data; the way the data is being processed; the impact on (in particular) fundamental rights (to freedom and non-discrimination) and further interests of the data subjects concerned; as well as the individual's' "reasonable expectations".[125]

2.2.2.2.1 Context of data collection

Regarding the context of the data collection, it is possible to differentiate, for example, between the spheres or media that are explicitly covered by fundamental rights. Therefore, one can differentiate between the collection of data that intrudes an individual's private home, or intercepts his or her communication, or takes place in the public. It then depends on those specific guarantees, how an individual's interest should be respected.[126]

The German Constitutional Court denies, for example, an intrusion into the scope of protection of the basic right of telecommunications (Art. 10 GG) only if the data is unintentionally collected — so only for technical reasons — and not immediately as well as irreversibly deleted after it has been collected.[127] In contrast, regarding the data collection in the public — irrespective of an individual's home or communications — the Court considered a smaller scope of protection. In the case of an automated registration of license plates, the Court only affirmed an intrusion into the scope of the right to informational self-determination if the collected data has been matched with certain keywords *and* leads to a positive result. Thus, the difference in the scope of protection of the right of telecommunications was, at the time, that this right covers each matching even if this leads to a negative result. The substantial reason for this difference might have been that the right of telecommunications seeks to

---

[123] See above in section "II. 2.1.2 Discussion on sufficiently precise specification", referring to *Nolte*, Art. 89, cip. 36, in: Art.
[124] See with respect to the fact that "interests" actually include, from a legal point of view, legal rights and, therefore, do not have to be mentioned explicitly, Assion/Nolte/Veil, Art. 6 cip. 131, in:Schlender/Stentzel/Veil Kommentar zur DS-GVO, 2nd ed., 2020; *Albers*, in: Wolff/Brink BeckOK DatenschutzR, 31. Ed., 2019, Art. 6 cip. 50.
[125] See Art. 29 Data Protection Working Party, Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC, 9 April 2014, 844/14/EN, WP 217, pp. 37 to 40.
[126] For example, cf. ECtHR, Case of Peck vs. the UK, 28 January 2003 (application no. 44647/98), cip. 61: "In those cases, the Commission attached importance to whether the photographs amounted to an intrusion into the applicant's privacy (as, for instance, by entering and taking photographs in a person's home), whether the photograph related to private or public matters and whether the material thus obtained was envisaged for a limited use or was likely to be made available to the general public. In (../the second case) the Commission noted that there was no such intrusion into the 'inner circle' of the applicant's private life, that the photographs taken of a public demonstration related to a public event and that they had been used solely as an aid to policing the demonstration on the relevant day. In this context, the Commission attached weight to the fact that the photographs taken remained anonymous in that no names were noted down, the personal data recorded and photographs taken were not entered into a data-processing system and no action had been taken to identify the persons photographed on that occasion by means of data processing (ibid.). Similarly, in (../the first case), the Commission specifically noted that the police used the photographs to identify offenders in criminal proceedings only and that there was no suggestion that the photographs had been made available to the general public or would be used for any other purpose."
[127] See BVerfG, 14th of July 1999, 1 BvR 2226/94 (Surveillance of Telecommunications) = NJW 2000, 55, cip. 160.

"avoid that the exchange of opinions and information through means of communications systems stops the data subjects concerned from communicating with each other because they fear that state institutions will access the content of the communication."[128] In the public domain, however, the German right to informational self-determination protects the individual against the risk that results from the storage of data providing the basis for potentially further measures.[129] Only in the case of a positive match, the Court affirmed in that decision such a risk because "[F]from this point in time, the license plate recorded is available for the processing by state agencies and the specific danger for the freedom of action and of being private occurs, which justifies the protection of the basic right to informational self-determination."[130] Thus, if the collection of data occurs in the public, it does not mean that there is not protection at all but that there is less protection, or at least, another form of protection.[131] The German Constitutional Court cannot help, of course, interpret European fundamental rights. However, in light of its extensive judicature, it can serve as a source of inspiration how European fundamental rights *could* be interpreted.

In fact, concerning the subject matter of this assessment, the collection of data will not intrude into an individual's home, nor will it contradict the right of an individual to respect his communication as the data is collected independently of a communication process. This is likely to be even the case with respect to Wi-Fi data as long as the data collection takes place before the user connects with the system.[132] Rather, the data is collected in the public. Indeed, the Forum Privatheit discusses the question of whether data related to probe requests that are sent out by an individual's device can be considered as publicly available data. This poses an important question in regard to the current German legal situation because the German data protection laws privilege the processing of publicly available data in favour of the data controller. However, with respect to the Wi-Fi data, the Forum Privatheit disputes this question, as data collected in connection to probe requests were inappropriate and not deemed to be available for an undetermined group of recipients, which is, however, the definition of "publicly available data" stated by German law.[133] Regardless of whether this interpretation of German law is correct or not, this DPIA is for a different reason a dissenting opinion. This DPIA considers all types of data (i.e. by the CCTV cameras, the parking lot sensors, and the Wi-Fi system) as collected in public not only because the collection actually occurs in public, but rather because the created movement patterns can be observed in public from each other person. It would only be much more time-consuming to do it analogously. Thus, as long as only such behaviour is recorded that can be observed in public, this DPIA assigns such data collection to the public context. But again, this does not mean that such data is not protected. However, the interest of the subject may be considered to be lower if the data collection occurs in public than if it were to interfere, for instance, with an individual's right to respect his or her home.[134]

[128] See BVerfG, 14th of July 1999, 1 BvR 2226/94 (Surveillance of Telecommunications) = NJW 2000, 55, cip. 135.
[129] See BVerfG, 11th of March 2008, 1 BVR 2047/05 and 1 BvR 1254/07 (License Plate Recognition) = NJW 2008, 1505, cip. 68.
[130] See BVerfG, 11th of March 2008, 1 BVR 2047/05 and 1 BvR 1254/07 (License Plate Recognition) = NJW 2008, 1505, cip. 69: "Ab diesem Zeitpunkt steht das erfasste Kennzeichen zur Auswertung durch Staatliche Stellen zur Verfügung und es beginnt die spezifische Persönlichkeitsgefährdung für Verhaltensfreiheit und Privatheit, die den Schutz des Grundrechts auf informationelle Selbstbestimmung auslöst."
[131] See, regarding the lower intensity of an infringement by the collection of data in the public, for example, BVerfG, 11th of March 2008, 1 BVR 2047/05 and 1 BvR 1254/07 (License Plate Recognition) = NJW 2008, 1505, cip. 83.
[132] Cf. BverfG, decision from the 22nd of August 2006, 2 BvR 1345/03 (IMSI Catcher), cip. 55 to 62 = NJW 2007, 351 (353)
[133] See *Eisele* et al., Forum Privatheit, White Paper Privatheit in öffentlichen WLANs, 2017, pp. 56 and 57.
[134] Cf. Art. 29 Data Protection Working Party, Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC, 9 April 2014, 844/14/EN, WP 217, p. 39.

2.2.2.2.2 Nature of the collected (and further processed) data

In the case of "Digital Rights vs Ireland", the European Court of Justice upheld the legal fact that the fundamental rights to privacy and data protection under Art. 7 and 8 ECFR also protect data subjects against the collection of personal data in public. In this case, a directive had required telephone communication providers to retain traffic and location data. The European Court of Justice stated:

> The retained data included "data necessary to trace and identify the source of a communication and its destination, to identify the date, time, duration and type of a communication, to identify users' communication equipment, and to identify the location of mobile communication equipment, data which consist, inter alia, of the name and address of the subscriber or registered user, the calling telephone number, the number called and an IP address for Internet services. Those data make it possible, in particular, to know the identity of the person with whom a subscriber or registered user has communicated and by what means, and to identify the time of the communication as well as the place from which that communication took place. They also make it possible to know the frequency of the communications of the subscriber or registered user with certain persons during a given period."[135]

The Court drew the following conclusions, firstly from the findings and secondly from the corresponding substantial guarantee provided by the fundamental right to private life under Art. 7 ECFR:

> "(T)those data, taken as a whole, may allow very precise conclusions to be drawn concerning the private lives of the persons whose data has been retained, such as the habits of everyday life, permanent or temporary places of residence, daily or other movements, the activities carried out, the social relationships of those persons and the social environments frequented by them."[136] This led the Court to the conclusion that "(T)to establish the existence of an interference with the fundamental right to privacy, it does not matter whether the information on the private lives concerned is sensitive or whether the persons concerned have been inconvenienced in any way",[137] and therefore, that "so far as concerns the right to respect for private life, the protection of that fundamental right requires, according to the Court's settled case-law, in any event, that derogations and limitations in relation to the protection of personal data must apply only in so far as is strictly necessary (...)."

In fact, this decision referred to data collected by telephone communication providers, and not to providers of CCTV cameras, parking lot sensors or Wi-Fi access providers. However, in particular the data collected by the Wi-Fi system and even more so, the possible insights that can be gathered from the processing of that types of data are very similar, if not the same. Therefore, in summary, the right to private life under Art. 7 ECFR requires — be it directly or indirectly through secondary law — the data controllers of the intended system to limit the collection of personal data, i.e. the gathered insights about the private lives of the data subjects concerned to what is strictly necessary pursuant to the research purposes. Consequently, the more sensitive the insights are, the more stringent the necessity test becomes. In particular, taking the third data usage scenario into account, this may be the case if the parking lot sensors detect a traffic road law violation via the implemented dashboard and cockpit

[135] See ECJ C-293/12 and C-594/12 (Digital Rights vs Ireland") cip. 26.
[136] See ECJ C-293/12 and C-594/12 (Digital Rights vs Ireland") cip. 27.
[137] See ECJ C-293/12 and C-594/12 (Digital Rights vs Ireland") cip. 27.

functions, or similarly, if the CCTV cameras record a crime.[138] Therefore, if one has to expect to collect those kinds of information, the necessity test is especially strict. In summary, these data types cannot always disclose specific information about the data collected by the system, but it *can* reveal such information. This hypothetical conclusion results, ultimately, from the sheer volume of data collected.

2.2.2.2.3 How the data is being processed

The amount of the collected data leads us to the criterion of how the data is processed. In this regard, the Art. 29 Data Protection Working takes the following aspects into consideration:[139]

> "Assessing impact in a wider sense may involve considering whether the data are publicly disclosed or otherwise made accessible to a large number of persons, or whether large amounts of personal data are processed or combined with other data (e.g. in case of profiling, for commercial, law enforcement or other purposes). Seemingly innocuous data, when processed on a large scale and combined with other data may lead to inferences about more sensitive data (...). In addition to potentially leading to the processing of more sensitive data, such analysis may also lead to uncanny, unexpected, and sometimes also inaccurate predictions, for example, concerning the behaviour or personality of the data subjects concerned. Depending on the nature and impact of these predictions, this may be highly intrusive to the individual's privacy."

The intended data processing expands to a large amount of data; concerning movement patterns, these patterns can also lead to profiling if the hashes of the collected MAC addresses are not renewed at sufficiently short intervals. Finally, it is feasible to suggest that the collected data or at least the processing results could be made public. In fact, whether the data will be made public and if so, under what conditions, is not yet certain. This decision (or these conditions) will, therefore, play a key role in determining whether data processing for research and statistical purposes in the field of urban traffic management can be based on Art. 6 sect. 1 lit. f GDPR.

2.2.2.2.4 Impact on rights (and further interests) and "reasonable expectations"

Also, the situation about the effects on the rights (or other interests) of the persons concerned is worth discussing. Given that the data is processed for research and for statistical purposes in the field of urban traffic management, there is no negative effect *intended*, for example, on the rights to liberty or non-discrimination. However, there can be *unintended* effects on data subjects:

First, considering the thoughts by the Art. 29 Data Protection Working Party, it is possible that the loss of a person's location data can lead once it has come into the hands of criminals to a dangerous situation or even death for that person.[140] In this context, it must be emphasised, that processing data for research and statistical purposes in an urban traffic management environment does not necessarily imply these risks. Rather, these risks are an unwanted and abstract risk (the latter being the case because, in addition to the processing purpose, there are no other specific objective circumstances that could serve as a reason

[138] However, critics do not consider such data processing as falling under Art. 10 GDPR (Processing of personal data relating to criminal convictions and offences) because this Article protects, teleologically, against the stigmatising effect caused by a conviction of a public authority, thus, not against the legal consequences of having committed a crime, cp. *Gierschmann* in Schlender/Stentzel/Veil, Kommentar zur DS-GVO, Art. 10, cip. 24; *Plath*, in: Plath, BDSG/DSGVO, 3. Aufl., 2018, Art. 10 cip. 2-4; *Weichert* in Kühling/Buchner, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 10 cip. 5.

[139] Cf. Art. 29 Data Protection Working Party, Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC, 9 April 2014, 844/14/EN, WP 217, p. 39.

[140] Cf. Art. 29 Data Protection Working Party, Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC, 9 April 2014, 844/14/EN, WP 217, p. 38.

for assuming such a specific risk). In that sense, the question arises of how likely such an abstract risk is — and that, in fact, depends on further safeguards.

Secondly, the Forum Privatheit considers a negative impact on the "freedom of choice" of the respective data subjects. These authors take into consideration that these effects are partly because some people who are particularly "data protection sensitive" could avoid public spaces where data is collected because of the "extensive harm to their private lives".[141] With respect to this important aspect, it is helpful to distinguish between the reasons for such an impact. On the one hand, such an impact can result from actual harm to the data subjects and, on the other hand, from harm that potentially exists that those affected expect.[142] Actual harm can occur when the data is used in a certain way that reduces the individual's room for maneuver.[143] This can be the case when an unsuspicious person becomes the object of state investigations, which adds to their risk of being unreasonably suspected.[144] Another example is when the individual runs the risk of being stigmatised.[145] Also if the individual cannot defend him or herself against the informational measure, this can lead to direct harm to the individual.[146] Concerning the target of this DPIA, one must take the considerations of the previous paragraph into account. The data processing for research and statistical purposes in an urban traffic management environment do not cause *per se* actual harm to the research subjects. However, such actual harm can be caused by possible misuse of the data, i.e. for purposes other than research and statistics in an urban traffic environment. In this regard, the second and third data usage scenarios described above give an idea of how such potential abuse might look like. For example, if the data is used leading to an unjustified discrimination of data subjects or raising concerns about the fairness of legal procedures. Whether data subjects are effectively restricted in their freedom by such a later use of data depends in this context on safeguards that avoid or at least reduce such a non-specific risk.

In contrast to such actual harm, effects resulting from the expectations of the respective data subjects can also be discussed regarding their "reasonable expectations".[147] The issue surrounding "reasonable expectations" of the individual is complex for several reasons, such as because it is important to avoid that the data subjects suffer from the unspecific threat that someone else knows a lot about them, but are unable to know what exactly. This fear can lead to the situation that data subjects stop on their own exercising their fundamental rights to freedom.[148] This risk can become highly relevant with respect to

---

[141] See *Friedewald* et al., Forum Privatheit, White Paper Datenschutz-Folgenabschätzung. 2017, p. 36: "Problematisch ist es zudem, wenn datenschutzbewusste Nutzerinnen und Nutzer aufgrund der weitreichenden Eingriffe in die Privatheit durch WLAN-Technologien bestimmte Areale meiden und damit ihre Bewegungsfreiheit eingeschränkt wird."

[142] Compare, as a source of inspiration, the differences made by the German Constitutional Court regarding a real harm, BVerfG, 3rd of March 2004, 1 BvR 2378/98 (Big Eavesdropping Operation) = NJW 2004, 999, cip. 227; BVerfG, 4th of April 2006, 1 BvR 518/02 (Dragnet Investigation) = NJW 2006, 1939, cip. 103, and regarding a potential harm, BVerfG, 14th of July 1999, 1 BvR 2226/94 (Surveillance of Telecommunications) = NJW 2000, 55, cip. 207; BVerfG, 3rd of March 2004, 1 BvR 2378/98 (Big Eavesdropping Operation), cip. 230.

[143] Cf. *Britz*, G., "Informationelle Selbstbestimmung zwischen rechtswissenschaftlicher Grundsatzkritik und Beharren des Bundesverfassungsgerichts", in: W. Hoffmann-Riem (Hg.), Offene Rechtswissenschaft, Tübingen 2010, pp. 561-596. cip. 570 and 571; see regarding actual harm for further fundamental rights to freedom, Albers,in: Wolff/Brink BeckOK DatenschutzR, 31. Ed., 2019, Art. cip. 72.

[144] See BVerfG, 3rd of March 2004, 1 BvR 2378/98 (Big Eavesdropping Operation) = NJW 2004, 999, cip. 227; BVerfG, 4th of April 2006, 1 BvR 518/02 (Dragnet Investigation) = NJW 2006, 1939, cip. 103.

[145] See BVerfG, 4th of April 2006, 1 BvR 518/02 (Dragnet Investigation) = NJW 2006, 1939, cip. 106.

[146] See BVerfG, 13th June 2007, 1 BvR 1550/03 (Retrieval of Bank Account Master Data) = NJW 2007, 2464, cip. 111.

[147] Another term used, in this regard, refers to the so-called "chilling effects", see Staben, J., "Der Abschreckungseffekt auf die Grundrechtsausübung - Strukturen eines verfassungsrechtlichen Arguments", Tübingen: Mohr Siebeck, 2016, pp. 42-62; such chilling effects can also lead to harm the fundamental rights, such as the freedom of expression, see ECJ C-293/12 and C-594/12 (Digital Rights vs Ireland"), cip. 28, and BVerfG, 14th of July 1999, 1 BvR 2226/94 (Surveillance of Telecommunications) = NJW 2000, 55, cip. 207.

[148] See, for example, regarding the freedom of expression, ECJ C-293/12 and C-594/12 (Digital Rights vs Ireland), cip. 28, and BVerfG, 14th of July 1999, 1 BvR 2226/94 (Surveillance of Telecommunications), cip. 207.

the target of evaluation of this DPIA in light of the extensive data processing operations.

2.2.2.2.5 Interims conclusion

All these considerations lead us to ask whether these concerns need to be accepted by data subjects, or whether these are clearly factors that are detrimental to the application of Art. 6 sect. 1 lit. f GDPR? To answer this question, it should be emphasised that the "legitimate interests"-clause does not require that there is no risk at all, but that balance is needed. There may, therefore, be risks for the respective data subjects, but these risks must not override the data controllers as well as the public interest in the processing of data. The Art. 29 Data Protection Working Party highlights this consideration in the following:

> "Finally, it is important to emphasise that not all negative impact on the data subjects 'weighs' equally on the balance. The purpose of the Article 7(f) balancing exercise is not to prevent any negative impact on the data subject. Rather, its purpose is to prevent disproportionate impact. This is a crucial difference."[149]

Taking these considerations into account, the question thus is whether additional safeguards could be implemented in such a way as to reduce the risks to data subjects to a level at which those risks (or interests) do not override the interests of the controllers (as well as third parties) and the public.

2.2.2.3 Balancing exercise taking additional safeguards into account

Additional safety measures implemented can hence "help 'tip' the balance on the scale" when carrying out the balancing exercise.[150] The next sections will illustrate at a general level what additional safeguards could be implemented to reduce previously reported risks in the following categories: first, the context of data collection and the nature of the data; second, the way in which the data is processed and the potential impact on the respective data subjects; and third, their "reasonable expectations". The chapter concludes with protection measures enabling data subjects to control still remaining risks, on their own.

2.2.2.3.1 Data minimisation according to the context and nature of the data

Concerning the context of data collection, the data is collected in public and therefore is less intensive in this regard than if the data were collected by intercepting a person's communication or intruding their home. However, the collection of personal data in a public context is also covered by the GDPR. In this regard, two aspects are particularly relevant: First, additional safeguards must address the risk that the collected data is processed in a way that serves as a basis for additional measures — be it by public agencies or private companies — that negatively affect an individual. Second, additional safeguards must defuse insights that may reveal the processing of data collected in the public domain on the private lives of data subjects (e.g., their patterns of movement). Data minimisation techniques (e.g. anonymisation and pseudonymisation) can play an important role in both risk mitigation strategies.

Addressing both risks, i.e. that the collected data reveals insights about the private lives of data subjects

[149] See Art. 29 Data Protection Working Party, Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC, 9 April 2014, 844/14/EN, WP 217, URL: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp217_en.pdf, p. 41.
[150] See Art. 29 Data Protection Working Party, Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC, 9 April 2014, 844/14/EN, WP 217, URL: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp217_en.pdf, p. 42.

and/or serves as a basis for possible further measures against them, the following aspects are particularly relevant:

- First, as a provider of CCTV cameras, KiwiSecurity should not only obscure the visual recordings so that people and vehicles can no longer be detected (especially faces and license plates). Rather, the raw material for identifying data subjects and/or vehicles should also be deleted immediately and irreversibly after it is collected.

- Second, Cleverciti, as the provider of the parking lot sensors, should avoid the situation where third parties, such as public order agencies, may mark certain areas as “forbidden zones” to use this information for law enforcement purposes. This means that these parties cannot use the system to monitor whether people park their vehicles in such "forbidden zones" and cannot relate that information to data subjects through additional information, such as the information obtained on the actual site. (The same situation should be avoided regarding the anonymised CCTV camera data.)

- Third, Cisco, as the provider of the Wi-Fi network, should hash the MAC addresses as described above (i.e. by adding a SALT). These hashes should be renewed frequently, both after a certain time interval and in relation to specific areas. This means that an individual cannot be tracked using a hash-value that serves as a unique identifier beyond that particular period and after the person has left that specific space.

Regarding the last aspect, the time intervals and the extent of the areas must be determined (e.g. one month, one week, or one day; and the radius of ten or five or one kilometre). This determination should result from a strike balancing exercise: On the one hand, one must determine the risks to the data subjects. This risk has two dimensions so far: First, the time interval and the spatial extent determine the insights that can be gathered about the private life of an individual moving in a particular space during a given period. Second, the time interval and the spatial extent determine the risk that the generated motion pattern per se serves as a unique identifier; and / or that the mere information that someone (albeit previously undiscovered) was in a particular place at a particular time can be used against a person (because that information is later used against him or her) The shorter the time interval and the smaller the area, the lower this risk is for the individual. On the other hand, the time interval must be so long and the area large enough that the collected data is still useful for research and statistical purposes.

2.2.2.3.2 Controlling the extent and potential impact of the data processing

Given the remarkable extent and impact of data processing on data subjects, it should again be underlined that the processing of research and statistics in an urban traffic management environment does not deliberately lead to adverse effects on the respective data subjects. However, in order to reduce the risk of unwanted negative effects on the data subjects that could be caused by the later data, it is necessary to control, in addition to the data minimisation of the data as described above, how the data is accessed and used later. This risk is at stake because of the many ways to combine the data collected from both the sources in the system and other external information. The last aspect is mainly the case when the data is made public without restrictions on such a combination. This in turn leads to information that is almost unpredictable. To minimise the risks of such unpredictable data connectivity, two main conditions should be met:

- First, even if all the data itself is at least pseudonymised, the combination of the collected data per se (that is, independent of other information collected outside the system) should only be

allowed if there are other control mechanisms; these mechanisms must control the risk that specifically results from the unpredictability of the information gathered through the data combination.[151]

– Second, and more importantly, the combination of that data with other information (coming from sources outside the system) must be controlled; this does not mean, that the data must not be made public at all, but rather there must be at least one mechanism that controls the conditions under which further information can be added.[152]

As emphasised before, this assessment concludes that the research and statistical purposes per se do not constitute a specific risk against an individual's fundamental right, such as the rights to freedom and/or non-discrimination (or any further interests). In principle, however, the data can always be misused in one way or another. Therefore, for this reason as well, the access and use of that data must be controlled (in a non-discriminatory way).

2.2.2.3.3 Transparency measures framing "reasonable expectations"

Last but not least, the measures described above as such do not avoid the situation that the data processing contradicts the "reasonable expectations" of the respective data subjects. To avoid this situation, the following aspects must be made as transparent as possible:

– Data processing
  - Different types of collected data (i.e. by the CCTV cameras, parking lot sensors, and Wi-Fi network)
  - Purpose(s) of data collection (i.e. providing technical services as well as research and statistics for urban traffic management)
  - How the data is processed (i.e. where the data is stored and processed and which entity has access to which data)

– Mechanisms of control
  - Data minimisation (i.e. of collected CCTV camera data, and Wi-Fi network data, including time intervals and spatial limits regarding renewals of hash-values)
  - Linkability and usage controls (i.e. regarding various data sources in the system and additional information from outside)

– Remaining risks
  - Insights into private lives based on collected movement patterns
  - Movement patterns as unique identifiers (combined with additional information, e.g. retrieved locally)
  - Relation of pseudonymous data to data subjects by means of additional information (e.g.

[151] Cf. *Elliot* et al. "The Data Anonymisation Decision-Making Framework" Ukan publications, 2016, pp. 52 seq and 90 seq.
[152] Cf. *Elliot* et al. "The Data Anonymisation Decision-Making Framework" Ukan publications, 2016, pp. 108 seq.

recovered locally)

To make these aspects transparent for the data subjects, as mentioned before, has the main objective that the data subjects do not stop exercising their fundamental rights to freedom (e.g. crossing areas where the data is collected) because they suffer from the unspecific threat that someone else knows a lot about them, but are unable to know what exactly and what the other one is going to do with that information.[153]

2.2.2.3.4 Opting-out as a form of managing remaining risks

One important measure for data subjects who want to object the data processing is to opt-out. This can be particularly important for data subjects who want to control the risks that still remain despite all the aforementioned measures. In this regard, the Forum Privatheit outlines the risk to a person's "freedom of choice" not only if that person stopped passing certain public spaces,[154] but also because they may not be able to decide on the collection of data that relates to them. In particular, these authors take into account situations where people are not asked for their consent at all, or that they are asked, but the free use of public Wi-Fi is made on the condition that the person gives their consent. Such an inappropriate restriction on the person's "freedom of choice" may be the case, in particular, where the person had to consent to the analysis of their online behaviour, the combination of the collected data with other data sets or the disclosure of the data to third parties.[155] These practices can lead, indeed, to excessive risk for the data subjects involved. In this context, however, it should be made clear that the lack of consent does not per se constitute an undue risk, but only if there are no other guarantees that sufficiently control the risks for those concerned.[156] With respect to the processing of personal data for research and statistical purposes in the urban traffic management environment, this DPIA concludes that this processing may be based on the "legitimate interest"- clause, independently of the individual's consent, if additional safeguards are fulfilled. The previous sections have demonstrated how these additional safeguards can adequately control the risks to the data subjects concerned. However, an additional safeguard should also entail to make the possibility to opt-out from the data processing as easy as possible for the respective data subjects concerned. Such an opt-out mechanism enables data subjects to even exclude risks that still remain after the other measures have been applied.[157]

The Art. 29 Data Protection Working Party considers such a safeguard as:

> "Additional measures may include, for example, providing an easily workable and accessible mechanism to ensure an unconditional possibility for data subjects to opt-out of the processing. These additional measures may in some (but not all) cases help tip the balance and help ensure that the processing can be based on Article 7(f), while at the same time, also protecting the rights and interests of the data subjects."[158]

It is sometimes argued that data subjects could opt-out by switching the Wi-Fi client of their devices off.[159] However, disabling the Wi-Fi client of a device is not an "easily workable and accessible"

---

[153] See, for example, regarding the freedom of expression, ECJ C-293/12 and C-594/12 (Digital Rights vs Ireland), cip. 28, and BVerfG, 14th of July 1999, 1 BvR 2226/94 (Surveillance of Telecommunications = NJW 2000, 55,), cip. 207.
[154] See above in section "II, 2.2.2.2.4 Impact on rights (and further interests) and 'reasonable expectations'".
[155] See *Friedewald* et al., Forum Privatheit, White Paper Datenschutz-Folgenabschätzung. 2016, p. 36.
[156] Cf. regarding the new right to data protection under Art. 8 ECFR, *Kranenborg*, H. in: Peers/Hervey/Kenner/Ward (eds.), The EU Charter of Fundamental Rights, A Commentary, Oxford i.a.: Hart 2014 Art. 8 (Protection of Personal Data) cip. 8.176.
[157] Cf. *Buchner*, B.,"Informationelle Selbstbestimmung im Privatrecht", Tübingen: Mohr Siebeck, 2006, p. 234.
[158] See Art. 29 Data Protection Working Party, Opinion 6/2014 on the notion of legitimate interest of the data controller under Article 7 of Directive 95/46/EC, 9 April 2014, 844/14/EN, WP 217, p. 41.
[159] See http://money.cnn.com/2011/11/22/technology/malls_track_cell_phones_black_friday/index.htm.

mechanism, as this would mean that the data subjects concerned could no longer use the Wi-Fi client of their devices. With respect to the current Smart City research project, such an "easily workable and accessible" opt-out mechanism could instead be implemented by offering data subjects the possibility to indicate on a website (e.g. provided for by the Wi-Fi network provider) that they do not want to be tracked. This would require data subjects to enter their MAC address on the specific website so that the network provider is able to immediately delete all collected data in relation to this address after a Wi-Fi access point has captured the data.[160] The information about this opt-out mechanism could be provided in two ways: First, on a general level, by reaching all those who enter (possibly) public spaces in which MAC addresses are being tracked. Second, on a more concrete level, by installing physical contact points in an area where the hash-value of a particular MAC address remains the same,[161] informing those data subjects who specifically pass the contact point. On such a website, the data subjects concerned could not only opt-out from the tracking but also, conversely, specifically agree to a certain data processing that could not be based on the "legitimate interests"-clause. Such an additional opt-in mechanism indeed raises the question of how such a mechanism should be provided that enables the data subjects to effectively manage their data protection risks. An answer to these questions, however, shall be discussed in later works that can build on this DPIA.

2.2.3 Interims conclusion

All these conditions are (still) somewhat generic. However, in subsequent DPIAs (or similar data protection risk assessments), which may build upon this legal-scientific DPIA, these conditions may be further specified. Assuming that all conditions are met, this DPIA holds it possible that the interests of the respective data subjects do not exceed the interests in the data processing. Therefore, the data processing for research and statistical purposes in the field of urban traffic management could be based, in principle, on Art. 6 sect. 1 lit. f GDPR.[162]

2.3 Limitation of the data processing and storage

The last step of assessing whether the intended data processing is necessary and proportionate refers to the requirements of purpose and storage limitation. Both requirements may be a significant burden on research and processing activities, particularly in relation to research and statistics. The reason for this is that research and statistics are often characterised by their open-ended processes. Future research purposes can thus hardly be predetermined at the moment of collection.[163] This fact typically conflicts

---

[160] Cf. the cooperation of several US-based companies offering such an opt-out solution, available at https://optout.smart-places.org/.

[161] See above in section "II, 2.2.2.3.1 Data minimisation according to the context and nature of the data"

[162] Cf. Art. 29 Data Protection Working Party, Opinion 12/2011 on smart metering, 4 April 2011, 00671/11/EN, WP 183, p. 14: "The key point to be made here is that reliance on this legal basis depends on giving proper weight to the interests and rights of data subjects. It might seem inarguable that the legitimate interests of the data controller and society as a whole would be served by increased efficiency in energy supply and consumption and that this might be achieved via the personal data collected from smart meters. However, simply because this particular use of personal data seems legitimate (and, to many people, desirable) does not mean that it can be applied to legitimise every element of processing. In other words, the imperative to reduce energy consumption, although it might be a sensible public policy objective, does not override data subjects' rights and interests in every case. Indeed, it is clear that including practical measures such as Privacy Enhancing Technologies and Privacy Impact Assessments to enhance the security and privacy of the data processed by smart meters will make it more likely that this condition for processing could be available to a data controller. This is particularly important where processing for a data controller's legitimate interests is both inherently and disproportionately intrusive or where the effect of the processing is to cause unwarranted detriment to the data subject. Examples might include the creation of detailed profiles of data subjects that are, in fact, not needed to achieve the purpose, passing details to third parties without the knowledge or consent of the data subject, or the use of personal data to take decisions about remote disconnection without proper regard for an individual's data protection and other rights."

[163] Cf. *Schumpeter*, Capitalism, Socialism and Democracy, pp. 82 and 83; Drucker, The Discipline of Innovation, in: Harvard business review 80 (2002): 95-104; *Fueglistaller* et al., Entrepreneurship: Modelle-Umsetzung-Perspektiven Mit Fallbeispielen

with the requirements that data must not be, at least principally, processed for other purposes than originally specified, and be deleted if it is no longer necessary anymore to achieve these original purposes.[164] Respecting the needs of research and statistics, the legislator has therefore established certain legal privileges regarding both requirements. Both the standard requirements and the legal privileges are going to be assessed in the two following sections. It shall be emphasised, in this context, that these privileges would not apply as far as the Proposal of the ePrivacy Regulation applied.

2.3.1 Purpose limitation (Art. 5 sect. 1 lit. b) GDPR)

With respect to the data collected by the system, the requirement of purpose limitation under Art. 5 sect. 1 lit. b GDPR does not conflict with the intended processing activities. Pursuant to Art. 5 sect. 1 half sent. 1 GDPR, personal data must not be further processed in a manner that is incompatible with the purposes that were originally specified. The requirement, thus, implies that the later data processing has a purpose other than the purpose of the collection. As long as the collected data for research and statistical purposes in the field of urban traffic management are processed for exactly this purpose, the purpose limitation requirement does not conflict.

However, the requirement may cause a conflict at the moment where the collected data is supposed to be combined with personal data collected outside of the system. This might be the case, for example, if certain analysis results regarding traffic jams are correlated with the sentiment analysis of personal data collected in social networks. In this case, social networks most likely did not collect personal data for the described research and statistical purposes, but for other purposes. Thus, pursuant to the requirement of purpose limitation, this personal data could be combined with data collected by the system only if this is not incompatible with the original purpose(s) (e.g. of the social networks). Regardless of the question on under which conditions a new purpose has to be considered as compatible or incompatible, this requirement fundamentally conflicts with the desired research processes.

However, in this regard, Art. 5 sect. 1 lit. b) half sent. 2 GDPR provides a legal privilege stating that further processing for scientific research purposes or statistical purposes "shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes". Pursuant to this legal presumption, personal data can therefore be further processed, regardless of the original purpose of collection given that this occurs for scientific research or statistical purposes.

2.3.2 Storage limitation (Art. 5 sect. 1 lit. e) GDPR)

The requirement of storage limitation is similar to the requirement of purpose limitation. Pursuant to Art. 5 sect. 1 lit. e) GDPR, personal data must not be "kept in a form which permits identification of data subjects for (..) longer than [it] is necessary for the purposes for which the personal data are processed". This means that all data must be de-identified at the moment when the research and statistical purposes in the area of urban traffic management do not require the identification of data subjects anymore. This requirement can principally conflict with the research goals pursued by participants of the system because they can hardly pre-determine — due to the (typically) open-ended research processes — if an identification of data subjects might be necessary at the exact moment of the assessment or at a later stage.

---

aus Deutschland, Österreich und der Schweiz. Springer-Verlag, 2012. Chapter: Entrepreneurship – Innovation and Entrepreneurship, p. 98; Moroz and Hindle. Entrepreneurship as a process: Toward harmonizing multiple perspectives. in: Entrepreneurship Theory and Practice 36.4 (2012): 781-818

[164] See Forgó et al., "The principle of purpose limitation and big data." New technology, big data and the law. Springer, Singapore, 2017. 17-42.

However, in this context too, the law provides for a legal privilege for scientific research and statistical purposes as Art. 5 sect. 1 lit. e) half sent. 2 GDPR states: "personal data may be stored for longer periods insofar as the personal data will be processed solely for (...) scientific (...) research purposes or statistical purposes in accordance with Article 89(1)" (emphasis added). Given this provision, the data collected by the intended system can, therefore, be stored longer, even if at a certain moment it is still unclear whether the data must later be associated with an identified or identifiable person to achieve the research objective. In any case, the requirement that the data must be processed *solely* for scientific research and/or statistical purposes makes it clear that the data must be separated from the environment in which it was processed for other reasons than research or statistical purposes.[165]

2.3.3 Legal privileges for research and statistical purposes (Art. 89 GDPR)

In fact, both legal privileges apply only if the two legal preconditions are met: First, the processing activities pursue "research and statistical purposes" within the meaning of the law; and second, the safeguards required under Art. 89 sect. 1 GDPR are fulfilled.

2.3.3.1 Definition of "scientific research" and "statistical" purposes

The law itself does not provide for a legal definition of "research" or "statistical" purposes. With respect to "scientific research" purposes, at least, "science" can be understood, in the style of a definition of the German Constitutional Court, as "the trial to gather, based on a certain state of knowledge, true (new) knowledge by thinking and working in a methodically organised, critical-reflective and discursive manner" (word in brackets added by the author).[166] In this regard, it is important to note that the European legislator has not chosen the notion of "scientific purposes", as often used under the Data Protection Directive, but of "scientific *research*" (emphasis added). Pursuant to some authors, the legislator sought, by choosing this wording, to react to new developments of Big Data and Data Mining restricting the legal privileges to "scientific research" in a narrow meaning.[167] However, recital 159 GDPR makes clear that "the processing of personal data for scientific research purposes should be interpreted in a broad manner including for example technological development and demonstration, fundamental research, applied research and privately funded research". Also private bodies can hence be considered as conducting "scientific research", even if these bodies *also* pursue economic goals.[168] To sum up, as long as the data processing pursues scientific *research* purposes, and not other scientific activities such as administration, the term can be broadly interpreted, including applied research funded by private parties, which may be the case in the future system.

Pursuant to recital 162 sent. 3 and 5 GDPR, "(S)statistical purposes mean any operation of collection and the processing of personal data necessary for statistical surveys or for the production of statistical results. (...) The statistical purpose implies that the result of processing for statistical purposes is not personal

[165] See *Herbst* in Kühling/Buchner, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 5, cip. 69-71, *Plath* in: Plath BDSG/DSGVO, 3. Aufl., 2018, Art. 5 Rn. 14; with regard to the nature of the term "scientific research purposes" and its distinction towards other forms of data controlling, see *Buchner/Tinnefeldt*, in: Kühling/Buchner, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 89, cip. 12 and Albrecht /Jotzo, Das neue Datenschutzrecht der EU: Grundlagen, Gesetzgebungsverfahren, Synopse. Baden-Baden, Germany: Nomos, 2017, part 3 cip. 71.

[166] See Nolte, Art. 89 cip. 18, in: Schlender/Stentzel/Veil Kommentar zur DS-GVO, 2nd ed., 2020, referring to BVerfGE 35, 79/113; 47, 327/267; 90, 1/12; *Buchner/Tinnefeld*, in: Kühling/ Buchner DS-GVO, Art. 89 cip. 12-13; referring to recital 159 GDPR Pauly also opines for a broad sense of the term: *Pauly*, in: Paal/Pauly, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 89 cip. 7

[167] See Albrecht/Jotzo, Das neue Datenschutzrecht der EU: Grundlagen, Gesetzgebungsverfahren, Synopse. Baden-Baden, Germany: Nomos, 2017, part 3 cip. 71.

[168] See *Nolte*, Art. 89 cip. 21 and 48, in: Schlender/Stentzel/Veil Kommentar zur DS-GVO, 2nd ed., 2020; in this regard, the only condition is that the legal privileges for "scientific research" purposes must not be transferred to these other economic purposes (see Art. 89 sect. 4 GDPR); similar *Pauly* in: Paal/Pauly, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 89 Sec. 4 cip. 18.

data, but aggregate data, and that this result or the personal data are not used in support of measures or decisions regarding any particular natural person." On the one hand, the notion of "statistic" purposes thus is narrower than "scientific research" purposes because a processing activity pursuing the identification of an individual does, *per definitionem*, not occur for a "statistic" purpose. In contrast, the processing of data for "scientific research" purposes only requires to draw *general* conclusions, but do not forbid, *per se*, the identification of data subjects.[169] On the other hand, the notion of "statistical" purposes is also wider than that of "scientific research" because statistics can also be done for pure private or economic matters.[170] However, sent. 4 of recital 162 GDPR, clarifies that "statistical results may further be used for different purposes, including a scientific research purpose." Thus, "statistical" and "scientific research" purposes can indeed overlap.

2.3.3.2 Safeguards required under Art. 89 sect. 1 GDPR

Given that the data processing activities pursue scientific research and/or statistical purposes, the legal privileges only apply if the further conditions under Art. 89 sect. 1 GDPR are met. Art. 89 sect. 1 GDPR states as: "Processing for (...) scientific (...) research purposes or statistical purposes, shall be subject to appropriate safeguards, in accordance with this Regulation, for the rights and freedoms of the data subject. Those safeguards shall ensure that technical and organisational measures are in place in particular in order to ensure respect for the principle of data minimisation. Those measures may include pseudonymisation provided that those purposes can be fulfilled in that manner. Where those purposes can be fulfilled by further processing which does not permit or no longer permits the identification of data subjects, those purposes shall be fulfilled in that manner." Keeping other provisions of the GDPR in mind, Art. 89 sect. 1 does actually not add further requirements to the processing of personal data. The reason for this is that the required safeguards (i.e. purpose minimisation, pseudonymisation and anonymisation) are already required under provisions applying to any kind of data processing (such as Art. 5 sect. 1 lit. b), c) and e), as well as Art. 25 and 32 GDPR). Insofar, Art. 89 sect. 1 GDPR hence rather emphasizes the importance of these requirements in particular for data processing activities taking place in the context of scientific research and statistics.[171] However, if the personal data is originally collected (also) for another purpose than scientific research or statistical purposes, the data used for scientific research and/or statistical purposes must be stored in a separate environment. The reason for this is that the legal privileges for scientific research and statistical purposes only apply, as described, to these purposes, and must not be transferred to other purposes, for instance, by using the same factual processing environment.[172]

2.3.3.3 Appropriate solution for open-ended research processes: From fully centralised (not so likely to be compliant) solutions to decentralised (more likely compliant) solutions

The legal privileges regarding the requirements of purpose limitation and storage limitation essentially offer an adequate solution for open-ended research processes, such as in an urban traffic management

[169] Cp. *Nolte*, Art. 89 cip. 20 and 22, in: Schlender/Stentzel/Veil Kommentar zur DS-GVO, 2nd ed., 2020; *Raum* in: Ehmann/Selmayr Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 89 cip. 89.
[170] See Nolte, Art. 89 cip. 22, in: Schlender/Stentzel/Veil Kommentar zur DS-GVO, 2nd ed., 2020; *Grages*, in: Plath BDSG/DSGVO Art. 89 cip. 6.
[171] See *Nolte*, in: Schlender/Stentzel/Veil, Kommentar zur DS-GVO, Art. 89, cip. 25; *Raum*, in: Ehmann/Selmayr Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 89 cip. 22 and 34-37; *Pauly* in: Paal/Pauly, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 89 cip. 15, *Caspar*, in: Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht, Art.89 cip. 49.
[172] See *Buchner/Tinnefeldt*, in: Kühling/Buchner, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 89, cip. 12 and Albrecht/Jotzo, Das neue Datenschutzrecht der EU - Grundlagen, Gesetzgebungsverfahren, Synopse, Nomos, Baden-Baden: Nomos p. 81 cip. 71.

environment. Applying the requirements of purpose limitation and storage limitation strictly, would instead conflict with such an openness of research processes. It must be stressed, however, that the legal privileges are only justified if appropriate safeguards are applied to protect the rights and freedoms of the data subjects concerned. In this regard, this legal-scientific DPIA provides for a basic structure that may serve – comparable to the safeguards that have been discussed in relation to Art. 6 sect. 1 lit. f) GDPR – as a basis for a more detailed risk assessment carried out later on. The balancing exercise in relation to Art. 6 sect. 1 lit. f) GDPR has already provided for certain guidelines with particular respect to data minimisation (by anonymisation and pseudonymisation), access and usage control. Different degrees of decentralization may be required either at a technical-infrastructure layer or rules layer (i.e. where involved stakeholders define the specific rules for the data processing) or both.[173]

Examples for such differentiating structures are provided, for instance, by Contreras and Reichmann with respect to scientific data pools. In their work, they observe the following "four basic structural models (...) along a continuum ranging from the most to the least centralized (see the table).

> (i) fully centralized: all data are aggregated in a single, centrally managed repository;
>
> (ii) intermediate distributed: repositories are distributed and separately maintained, but may be interconnected by a central access portal, share technical service components, and utilize a common data-exchange format [sometimes called a federated database system];
>
> (iii) fully distributed: repositories are maintained locally and are not technically integrated, but share a common legal and policy framework that allows access on uniform terms and conditions (legal interoperability);
>
> (iv) noncommons: repositories are largely disaggregated and lack technical and legal interoperability and, at most, may share a common index."[174]

Contreras and Reichmann also evaluate the advantages and disadvantages of these observed patterns from a mainly economic viewpoint: While both authors see the fully centralized models positively with respect to better data quality, but also negatively given its higher costs, they deter from the non-commons because of their complete lack of interoperability. As a consequence from this, both authors highlight the models in between that may provide for technical and/or legal interoperability but at lower costs than fully centralized models.[175]

These examples make it clear that the collected data for the technical provision of the services in the intended system do not have to be stored at all, not even in a separate environment, for the research and statistical purposes. Instead, it is possible to provide just a common legal and policy framework for the case that the data shall be processed (and eventually combined) for research and statistical purposes later on. For instance, real-time processing does not require storing the data in advance. However, a common legal and policy framework makes it much easier to combine and process the data for the research and statistical purposes in the moment when these purposes become acute. The reason for this is that the criteria and processes are already defined and implemented when these purposes become

[173] See in favour of decentralized structures, in general, iKoPA, Deliverable 3.2 – Datenschutz bei vernetzten, automatisierten und kooperativen Fahrzeugen nach der Datenschutzgrundverordnung, pp. 99 et seq., with further references, online available at: https://www.datenschutzzentrum.de/uploads/projekte/ikopa/iKoPA_D3.2-3.pdf; see in more detail in the research project "Data Governance", where the research group analyses different data governance models on a technical-infrastructure layer, an organisational layer, and a rule-making layer, online accessible under https://www.hiig.de/en/project/data-governance/.
[174] *Contrera/Reichman*, "Sharing by design: Data and decentralized commons." Science 350.6266 (2015): 1312-1314, p. 1312.
[175] *Contrera/Reichman*, "Sharing by design: Data and decentralized commons." Science 350.6266 (2015): 1312-1314, p. 1313.

relevant so that one does not have to start from scratch. Of course, in particular if the research and statistical purposes require the controller to store personal data over a longer period, the prepared structures must guarantee that the data are at least, pseudonymized. Against this background, a main question on the appropriate data protection by design-strategy for the intended system will therefore be which criteria and processes can be defined and implemented in advance as well as what criteria and processes can only be specified during the later research process when its purposes get more specific and acute.

3 Controller's duties and data subjects' rights

Given that data processing for scientific research and/or statistical purposes in an urban traffic management environment can be justified by the legitimate interests of the data controllers of the future system (and/or other third parties, and the public), it is necessary to assess whether, and if so, how the specific controllers' duties and data subjects' rights apply. In this regard, an essential question is whether the controller is actually able to identify and contact the individual. The reason for this is that the rights of the individual require the controller to relate the processed data to the person claiming those rights, and that the obligation to information basically implies contact with the individual. In contrast, as described above, the participants in the proposed system may only partially identify data subjects, if anything and there may be few, if at all, point of contact with them.[176] For these cases, the law provides for a special regulation.

3.1 Controller's duty of information (Art. 14 sect. 1 to 4 GDPR)

With respect to the controller's duty to inform the data subjects about its processing, the law differentiates between two situations: In the first situation the data has been directly collected from the data subjects concerned; in the second situation the data has not been collected from the data subjects, but indirectly (e.g. through another source). Despite this differentiation, a closer look reveals that the information is largely the same in both cases.[177] The two main differences concern, on the one hand, additional information that the controller must give the data subjects if the data is not directly collected from them and, on the other hand, two legal privileges for the controller over whether, when and how it must provide the information.

3.1.1 Information that has always to be provided

In principle, controller(s) of the proposed system must inform the data subjects in the moment the data is collected about the following aspects (section 1 of Articles 13 and 14 GDPR):

- The scientific research and statistical purposes in the area of urban traffic management for which the data is collected;
- who the controller(s) is (or are), and who obtains access to the data;
- the legal basis of Art. 6 sect. 1 lit. f) GDPR for the data processing, as well as the interests of the controller(s) and third parties (in terms of transparency, it is also useful to stress the public interests).

In addition, the controller(s) must inform the data subjects about the following aspects not only at the

[176] See the assessment above in section "II, Controls for urban traffic management - 1 Application of data protection laws – 1.2. Regarding CCTV data - 1.3. Wi-Fi data - 1.4 Regarding parking lot assistant data".
[177] Cf. Art. 13 and 14 GDPR.

time of collection but each time when they process the data for a different purpose (if it is "necessary to ensure fair and transparent processing", pursuant to section 2 combined with section 3 and 4 of Articles 13 and 14 GDPR):

- The exact length of time the data will be stored;
- the existence (if they may exist) of the rights to
  - access the data;
  - rectify the data;
  - delete the data;
  - object to the processing;
  - to file a complaint
- and, of course, always the new purpose(s).

3.1.2 Additional information and legal privileges if the data has not been obtained from the individual

Beside this, the law foresees two special rules if the data has not been directly obtained from the data subjects. In such cases, Art. 14 requires, on the one hand, the controller to additionally inform the data subjects about

- the categories of data collected in the moment the controller obtains the data;
- and both in the moment the controller obtains the data and when it processes the data for another purpose than for which it has obtained the data
  - the source from where the data originates;
  - and the interests of the controller(s) and third parties (as well as the public) but only if the processing is based on Art. 6 sect. 1 lit. f) GDPR.

On the other hand, Art. 14 GDPR essentially provides for two legal privileges in favour of the controller. One of these privileges is established under Art. 14 sect. 5 lit b) GDPR. According to this provision, the controller(s) of the proposed system will not have to provide the data subjects with the information, at least not directly, if the provision of the information is

- either impossible,
- or likely to render impossible or seriously impair the goal of the processing,
- or would involve a disproportionate effort.[178]

This may be particularly the case with respect to the intended system, that data often relates only

[178] See recital 62 sent. 3 GDPR stating, in this regard, that "the number of data subjects, the age of the data and any appropriate safeguards adopted should be taken into consideration."

indirectly to identifi*able* individuals.[179] In such cases, the controller is therefore hardly able to directly inform the data subjects.[180] However, even if Art. 14 sect. 5 lit. b) GDPR applied, sent. 2 requires the controller(s) to

> "take appropriate measures to protect the data subject's rights and freedoms and legitimate interests, including making the information publicly available".

Another legal privilege concerns the time at which the information must be provided. If the data is directly obtained from the data subjects, this must be the case before, or at least, in the moment of the collection, under Art. 13 sect. 1 sent. 1 GDPR). However, if the data is not obtained from the data subjects, the controller(s) must only give the information, based on Art. 14 sect. 3 GDPR,

- either within a reasonable period after the data has been collected (at the latest within one month) by taking the specific circumstances of the processing into account;
- if the data is used for communication with the person, at the latest at the time of the first communication with the person concerned;
- or if it is intended to be disclosed to another recipient, at the latest when the personal data are first disclosed.

3.1.3 Discussion on the collection of the data (directly from the individual or not)

Whether, how and when the controller must give which information, depend on the question of whether the data is directly obtained from the individual or not. However, among legal scholars, it is unclear how a distinction can be made between data obtained directly and data that was not obtained directly from the respective data subject. According to German data protection law, the essential criterion was and still is whether the individual knows that data is collected about him/her or not.[181] Some legal scholars also seem to apply this criterion in regards to the interpretation of Art. 13 and 14 GDPR.[182] For example, according to Schmidt-Wudy, the data is obtained directly from the data subjects, if they are involved in the collection of the data from the point of view of the controller (be it active or passive).[183] Similarly, Kühling/Martini consider Art. 13 GDPR only applicable if the data is directly collected from the individual openly and not secretly.[184] With regard to the future system, this opinion might lead to the result that the special rules under Art. 14 GDPR apply as long as the data subjects do not participate in, or at least, are not aware of the data collection.

In the intended system, it might be argued on the one hand that the data subjects participate in the collection of the data because they walk around in public spaces and/or park their vehicles in parking lots and/or have not turned off their Wi-Fi function of their personal devices. On the other hand, they may not know that the CCTV cameras, parking lot sensors and Wi-Fi access points that are installed collect the data for the purpose of urban traffic management. If this unawareness were to be sufficient

---

[179] See above in section "II, Controls for urban traffic management – 1.5. Interims conclusion".
[180] See Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007, 01248/07/EN, WP 136, p. 10-13; see also *Schmidt-Wudy*, in: Wolff/Brink BeckOK DatenschutzR Art. 14 Sec. 5 lit. b) cip. 98; see *Veil*, Art. 13 and 14 cip. 145, in: Schlender/Stentzel/Veil Kommentar zur DS-GVO, 2nd ed., 2020; *Paal*, in: Paal/Pauly, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 14 Sec. 5 lit. b) cip. 40.
[181] See *Kratschmer*, in: Gierschmann/Saeugling, Systematischer Praxiskommentar Datenschutzrecht, 2014, § 4, cip. 34.
[182] See, for instance, *Franck*, in: Gola, DS-GVO, Art. 13 cip. 4.
[183] See *Schmidt-Wudy*, in: Wolff/Brink BeckOK DatenschutzR, Art. 14 DS-GVO cip. 30 sequ.
[184] See *Kühling/Martini* in: Kühling/Martini et. al, Die Datenschutz-Grundverordnung und das nationale Recht. Erste Überlegungen zum innerstaatlichen Regelungsbedarf. Münster (2016). p. 406 sequ.

for the application of Art. 14 GDPR, the special rules would apply.

Against this opinion, some legal scholars argue that the European legislator has not chosen the awareness of the individual as the essential criterion but pursued a more objective approach, by referring to the "source" from where the data comes.[185] These authors outline that the subjective approach, referring to the knowledge of the individual, would lead to an incentive for the controller to collect the data secretly.[186] Such an incentive would be in conflict with the principle of transparency under Art. 5 sect. 1 lit. a) GDPR. Furthermore, this result would also lead to regulatory inconsistency since the individual's need for protection is actually higher if the data is not obtained directly from him or her than if the data is obtained directly.[187] Thus, data controller(s) of the future system would have an incentive to secretly collect the data in order to profit from the legal privileges under Art. 14 GDPR.

Nonetheless, looking at this second opinion more precisely, one can argue that there is actually no inconsistency within the law. The reason for this is that one can see both articles as the intended result of the balancing exercise of the legislator that has weighed the needs of the data subjects concerned against the needs of the controllers that are both embedded in certain processing environments.[188] Art. 14 sect. 5 lit. b) sent. 1 GDPR discharges, in a first step, the controller of its obligation to give the data subjects the information directly. However, in a second step, the second sentence requires the controller to take appropriate steps, such as making the information publicly available. This result does not necessarily lead to an incentive for the controller to secretly collect the data. If the requirement of public disclosure of the person is strictly interpreted, it is not necessarily easier for the controller to fulfil such an obligation to inform the public than to inform the individual directly. This may be the case, in particular when the public information needs to be so comprehensive that the risk that an individual might miss the information tends toward zero. This argument does not mean that this risk must be zero. Rather, this argument does only show that the legal privileges under Art. 14 GDPR do not necessarily lead to a regulatory inconsistency (if applied correctly).

Finally, it should be noted that the information duties do not apply if the person already has the information. This may be relevant to those participants of the intended system who do not collect the data themselves but use data that has already been collected. If the controllers collecting the data also inform the person about the processing purposes of these other participants (among other aspects, see above), these participants do not need to inform the data subjects again.

3.2 No data subjects' rights if the controller is not able to identify the data subject (Art. 11 and 12 sect. 2 GDPR)

With respect to the rights of the data subjects, Art. 11 and 12 sect. 2 GDPR explicitly address the situation where a data controller cannot relate the data to an *identified* individual (but the data subject is still *identifiable*).[189] In this case, the controller does not have to store (or collect more) information that identifies an individual only to be able to fulfill the specific rights and duties in Art. 15 to 22 GDPR.[190] Since the controller can avoid these rights and obligations by deleting such direct identifiers, both

[185] See recital 61 sent. 1 GDPR.
[186] Cf. the examples given by *Veil*, Art. 13 and 14 cip. 42, in: Schlender/Stentzel/Veil Kommentar zur DS-GVO, 2nd ed., 2020
[187] See *Veil*, in: Schlender/Stentzel/Veil, Kommentar zur DS-GVO, Art. 13 and 14, cip. 41.
[188] See *Grafenstein*, Art. 2 cip 1-6, in: Schlender/Stentzel/Veil Kommentar zur DS-GVO, 2nd ed., 2020
[189] Otherwise the GDPR would not apply at all.
[190] The interplay of both Articles 11 and 12 sect. 2 GDPR is rather confusing; for instance, while Art. 11 only refers to Art. 15 to 20 GDPR, Art. 12 sect. 2 GDPR refers to Art. 15 to 22 GDPR; in order to avoid this and further inconsistencies, the following assessment focuses on the principle idea of both provisions without going into too much detail; see for further information, for example, at *Veil*, Art. 11 cip. 54 et seq., in: Schlender/Stentzel/Veil Kommentar zur DS-GVO, 2nd ed., 2020; Paal in: Paal/Pauly, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 12 cip. 49-50; as well as *Plath*, in: Plath BDSG/ DSGVO Art. 11 cip. 1-9.

articles provide an incentive to de-identify the data.[191] If a controller cannot relate the data to an identified individual, the law provides for a three-step procedure: In a first step, the data controller has to "demonstrate that it is not in a position to identify" the individual. In a second step, the individual can now provide, on his or her own, further information enabling the controller to relate the data to him or her; this might be the case if the individual has a particular interest in the data processed by the controller (even if this data does not *directly* relate to him or her). In a third step, the controller must — being now able to relate the data to the individual — apply the specific rights and duties under Art. 15 to 22 GDPR. The following sections demonstrate this procedure in relation to the various types of data collected from the intended system.

3.2.1 CCTV camera data

The current DPIA has come to the conclusion that KiwiSecurity as the provider of the cameras, should pseudonymise, as far as possible, the data so that the processing of all data collected by the intended system for the scientific research and statistical purposes can be based on the "legitimate interests"-clause, pursuant to Art. 6 sect. 1 lit. f) GDPR. If the data was completely anonymised, the collected data could not even be related to an *identifiable* individual. In this case, the GDPR did not apply. With respect to Art. 11 and 12 sect. 2 GDPR, it is clear why: If the provider obfuscates the collected data (i.e. faces, vehicles, and license plates) sufficiently, the information given by an individual indicating that the person present in the video is him or her would be a blanket assertion. In this case, the individual has no need for his or her rights under Art. 15 to 22 GDPR. However, at the moment somebody can prove that this assertion is correct, the recorded information can be related to the recorded individual and, as a consequence, he or she has a need for the rights under Art. 15 to 22 GDPR. Many of these situations may refer to the execution or defence of legal claims. However, before getting into the details of such situations, the next section will illustrate, in light of the intended system, the technical complexity related to Art. 11 and 12 GDPR.

3.2.2 Wi-Fi access point data

As mentioned previously, the current DPIA has come to the conclusion that Cisco, as the provider of the Wi-Fi system should hash the MAC addresses immediately after it is collected. In this case, the provider adds a random number (called SALT) to these addresses and computes the hash based on this extended sequence. This hash allows the network provider to discover a particular device even if it has deleted the SALT and, therefore, cannot longer re-calculate the MAC address itself. However, to avoid the situation where the provider can analyse, based on a particular hash, a movement pattern of the device that allows re-identification of an individual, the current DPIA has additionally come to the conclusion that the hashes must be renewed, on a sufficiently frequent basis.[192] This procedure leads to the following conclusion regarding Art. 11 and 12 sect. 2 GDPR:

As long as the provider of the Wi-Fi system stores the SALT and, therefore, can recalculate the MAC address of a device, the provider is principally able to refer to the MAC address (and the data collected in relation to that address) to an identifiable individual. If the provider demonstrates that it is not in a position to identify, through the MAC address, the individual because further information is needed, this individual cannot claim the rights in Art. 15 to 22 GDPR. However, this individual can help the provider to make the connection by disclosing the MAC address of their device and authenticating them. This information enables the provider now to relate the location data associated to the hash to the

[191] See *Veil*, Art. 11 cip. 1, in: Schlender/Stentzel/Veil, Kommentar zur DS-GVO, 2nd ed., 2020; Plath, in: Plath, BDSG/DSGVO, 3. Aufl., 2018, Art. 11 cip. 1.
[192] See above in section "II, Controls for urban traffic management scenario – 2.2.2.2 Interests of data subjects".

device. This requires the provider to match the SALT stored within its system with the addresses given by the individual recalculating the hash for the person's device. In fact, as highlighted before, the provider can only transfer the hash back to the individual's device as long as it saves the related SALT. The moment the provider renews the hash and deletes the SALT, it can no longer refer the location data to the individual's device. This means that the person can only claim their rights under Art. 15 to 22 GDPR, as long as the provider has not yet renewed the hash and deleted the SALT. Thus, the rights only refer to data that has been collected for the period in which the hash is not renewed and the SALT is not yet been deleted.

The situation is similar from the perspective of an online service provider, which only gets the hash from the Wi-Fi provider. As long as the Wi-Fi provider stores the SALT, the hash and the data collected in relation to the hash can only be considered to be pseudonymised (but not anonymised). However, the online service provider cannot by itself relate, on its own, the hash to an individual. So far, the online service provider does not have to apply the rights and duties under Art. 12 to 22 GDPR if it demonstrates this situation toward the person who claims one or more of these rights. In a second step, the person can now disclose the MAC address of their device and authenticate themselves. This enables the online service provider to relate, *from now on*, all data collected in relation to the MAC address to the individual. For this data, the person can now claim their rights. However, with respect to the *data that has already been collected*, the online service provider additionally needs the SALT (from the Wi-Fi provider) in order to make the connection between this already collected data and the MAC address given by the data subject. Only if the online service provider has the SALT, it can recalculate the MAC address from the hash and relate the collected dala to the individual. At this moment, the person can thus claim their rights.

3.2.3 Parking lot sensor data

With regard to the data collected by the parking lot sensors, the situation is more complex. The structure of this assessment essentially follows the described data usage scenarios. As a first step, Cleverciti, as a provider of parking sensors, does not collect data related to an identified person. However, in a second step, the collected data relates to identifiable data subjects. This is the case in the second data usage scenario because an individual may want to connect with the Wi-Fi system in order to use the mobile application offered by Cleverciti helping this user to find available parking space. Once a person creates a personal account for this service, Cleverciti is able to relate the collected information about the person to them whilst they are using this service. This is certainly the case when the individual registers by authenticating him or herself, for instance, by giving his or her payment details, or any other information revealing his or her identity. In this case, the individual is *identified*, and the rights and duties under Art. 15 to 22 GDPR apply.

In contrast, if the person does not register through the authentication, he or she is only *identifiable*. As explained earlier, Cleverciti is able to identify the individual because it can at least relate the account (and the data collected while the user is logged on) to the MAC address of his or her device. The reason for this is that the MAC address is no longer hashed in the moment the person connects to the Wi-Fi system. Whether in this case Art. 11 and 12 sect. 2 GDPR apply or not, in a first step, depends on how the notion "unless the controller demonstrates that it is not in a position to identify the data subject" is interpreted. In this context, the question arises as to whether or not the controller's knowledge of a certain MAC address puts the controller "in a position to identify the data subject". However, instead of delving deeper into this first step of the legal analysis of Art. 11 and 12 sect. 2 GDPR, the current DPIA focuses on the second step, illustrating why an individual may have an interest in getting access to that data. This could be the case, for example, if the person wants to use the data to defend herself against the

claims of another person or entity. In such a case, it may be more straightforward for the individual to provide him or herself with the necessary information that will allow the controller to relate the data to the him or her (and grant him or her the access), rather than to demand from the provider to make the reference on its own (what may be more complicated and, thus, more cost and time consuming).

The third data usage scenario provides a vivid example for why a data subject may have an interest in gaining access to the collected data in order to defend him or herself against a legal claim. According to the first alternative of this scenario, the collected data relates to an *identifiable* individual because a public order agency could use the data for the enforcement of traffic law by relating the data to the individual on-site. After evaluating the legal basis for data processing for scientific research and statistical purposes, the current DPIA has concluded that Cleverciti should avoid the situation in which a public order agency is able to use the data for the purposes of law enforcement. However, similar to the case described above, it is imaginable that a public order agency might accuse the person to have committed a traffic law offense, irrespective of the collected data, but on the basis of other information, for example, because a representative of the agency was on-site and misinterpreted the situation. In such a case, the individual may seek access to the data collected by Cleverciti in order to prove that he or she has not committed the offense. Since in such a case, Cleverciti is certainly not able to relate the data to the individual, Art. 11 and 12 sect. 2 GDPR apply. The person must, therefore, provide the necessary information to prove that the data relates to them. If the individual does so, he or she can claim the rights under Art. 15 to 22 GDPR apply.

3.3 Further legal privileges for scientific and statistical purposes regarding the data subjects' rights (Art. 15 to 22 GDPR)

Even if an individual provides the information that enables the controller to link the processed data to the individual, the application of the rights under Art. 15 to 22 GDPR depend on further requirements. The reason for this is that the law foresees, here again, further legal privileges for the processing of personal data for scientific research and statistical purposes. In this regard, the regulation does not regulate all legal privileges. Instead, Art. 89 sect. 2 offers the EU Member States the opportunity to derogate these rights under national law (Art. 89 sect. 2 GDPR). This possibility of derogation concerns: the right to access (Art. 15 GDPR); the right to rectification (Art. 16 GDPR); the right to the restriction of the processing (Art. 18 GDPR); and the right to object the data processing (Art. 21 GDPR). The German legislator used this possibility by establishing § 27 sect. 2 BDSG_new. In contrast to such a derogation under national law, the GDPR itself regulates the application of the legal privilege with respect to the right to erasure (Art. 17 GDPR). Last but not least, it is important to note that the regulation does not only restrict these rights or allows its derogation for scientific research and statistical purposes, but also provides for a minor extension of the individual's legal position with respect to the right of objection (Art. 21 sect. 6 GDPR).

3.3.1 Rights to access, rectification, and restriction (Art. 15, 16 and 18 GDPR)

Pursuant to Art. 15 GDPR, an individual has "the right to obtain from the controller confirmation as to whether or not personal data concerning him or her are being processed, and, where that is the case, access to the personal data and the (...) information". This type of information essentially corresponds to that under Art. 13 and 14 GDPR.[193] Access to the data also serves the information basis for the data subject to rectify that data. In this regard, Art. 16 GDPR contains two with respect to its protection effects complementary rights: First, the data subject has "the right to obtain from the controller without undue delay the rectification of inaccurate personal data concerning him or her"; and taking into

[193] See, insofar, above in section "II, 3.1.3 Discussion on the collection of the data (directly from the individual or not)."

account the context where the information is used, the data subject has "the right to have incomplete personal data completed, including by means of providing a supplementary statement." Finally, a data subject may also request the controller to restrict the data processing, pursuant to Art. 18 GDPR. With respect to the intended system, this could be particularly relevant for data subjects, when they do not wish to erase the data but still need the data, for example, for "the establishment, exercise or defence of legal claims" (sect. 1 lit. c).[194] In any case, the person can request the deletion of the data only if this is no longer necessary for the achievement of the processing purposes.

This last aspect provides a good illustration of why the data subject may want access to the data relating to him or her and/or to correct the data. This may be the case, as shown before, since the data subject could use the data collected for example by the provider of the parking spot sensors in relation to legal claims.[195] In any case, the controller processing the data for scientific research and/or statistical purposes can also refuse such rights if its exercise is likely to render impossible or significantly affect the achievement of those purposes, pursuant to § 27 sect. 2 sent. 1 BDSG_new in combination with Art. 89 sect. 2 and 1 GDPR. In terms of access rights, the controller's room for denial is even a bit wider. In that regard, the controller may also refuse this right if its exercise would lead to disproportionate efforts for the controller, pursuant to § 27 sect. 2 sent. 2 BDSG_new in combination with Art. 89 sect. 2 and 1 GDPR.

3.3.2 Right to erasure and right to object the processing (Art. 17 and 21 GDPR)

If the data subject does not need the data anymore, he or she can also require, according to Art. 17 GDPR, the controller to delete the data if further requirements are met. With respect to the intended system, the following two cases may be particularly relevant. The data subject can require the deletion of data relating to him or her: first, if the data is no longer needed for the purposes for which they were collected (sect. 1 lit. a); and second, if the person objects to the processing and there are no compelling reasons for the processing (sect. 1 lit. c). While the first case does not place a special burden on the controller(s) of the future system, the second case could be more problematic. In the first case, individual rights do not apply as long as the data processing is still required for the scientific research and/or statistical purposes in the field of urban traffic management. In contrast, in the second case (i.e. if the individual objects to the processing), the controller must prove that it has valid reasons to deny the individual's claim to delete the data. This consideration is stricter than the one previously described with respect to Art. 6 sect. 1 lit. f) GDPR.[196] However, this second case becomes also less problematic as long as the deletion of the data "is likely to render impossible or seriously impair the achievement" of the scientific research and statistical purposes, according to Art. 17 sect. 3 lit. d) GDPR.

With regard to the data subject's right to object, the situation for the controller appears to be more stringent. Pursuant to Art. 21 sect. 6 GDPR, the right to object the data processing for research and statistical purposes applies if two cumulative conditions are met: first, the data subject objects the data processing "on grounds relating to his or her particular situation"; and second, the processing for research and statistical purposes is not necessary for a public interest task. Both components of this definition are interesting with respect to the future system: first, one may ask which reasons relating to a data subject's "particular situation" gives him or her the right to object?; and second, whether the data processing for research and statistical purposes in an urban traffic environment can be considered as a public interest task? However, while the second component would certainly be met, the German legislator made – also in this regard – use of its ability to restrict the data subjects' right to object.

[194] Cf. above in section "II, 3.2.3 Parking lot sensor data".
[195] Cf. above in section "II, 3.2.3 Parking lot sensor data".
[196] Cf. above in section "II, 2.2.2 'Legitimate interests'-clause (Art. 6 sect. 1 lit. f GDPR)".

According to § 27 sect. 2 sent. 1 BDSG_new (in combination with Art. 89 sect. 2 and 1 GDPR), the right to object does not apply, here again, if its exercise would make it impossible or significantly impair the achievement of the scientific research or statistical purposes. Therefore, the two components of Art. 21 sect. 6 GDPR become relevant only if the data subjects' right to object the processing did not seriously impair the research or statistical purposes.

3.3.3 "Likely to render impossible or seriously impair" the research purposes

As mentioned above, the GDPR itself laid down the conditions for applying the legal privileges concerning the right to object and the right to erasure. However, Art. 89 Sect. 2 GDPR does not explicitly refer to the right to erasure (Art. 17 Sec. 3 GDPR) when assigning the translation of legal derogations to EU Member States' national laws. Consequently, the German legislator did not include derogations in relation to Art. 17 Sec. 3 GDPR into § 27 BDSG_new.[197] Contrary to the individual's right to object laid down in Art. 21 GDPR, conditions for exemptions of the right to erasure are based on the minimum requirements of Art. 89 Sec. 1 GDPR. As such, Art. 17 Sec. 3 lit. d contains the aforementioned exclusion clause for the relevant right of erasure and comes into operation where it "is likely to render impossible or seriously impair the achievement of the objectives of the processing". Whether or not this is the case must be decided on a case-by-case basis. However, all privileges should be subject to the proviso that the processing is carried out in accordance with Art. 89 Sec. 1 GDPR, i.e. "guarantees for the rights and freedoms of the data subject under this Regulation". Accordingly, it refers to the general principles relating to the processing of personal data (Art. 5 GDPR), the lawfulness of processing (Art. 6 GDPR), the responsibility of the controller (Art. 24) and security of processing (Art. 32 GDPR).[193]

The extent to which claims based on Art. 17 Sect. 3 GDPR are feasible to jeopardise a scientific research project has not yet been sufficiently clarified. If the deletion of personal data might lead to an impairment of archiving purposes or of research matters, it must be evaluated on a case-by-case basis in form of a forecast.[198] At least, the impairment should not be negligible.[199] The provision may be applicable if the asserted rights are in some way likely to make data processing for research or statistical purposes more difficult, inefficient or even impossible, for example by imposing a heavy administrative burden or by depriving the bodies privileged by Art. 89 of their data and thus working basis. The achievement of the purposes of the processing may be affected, for example, in such a way that the absence of data significantly reduces the statistical validity of results, or in so far as the processing in question depends on the completeness of the data. However, the impairment must in any case be at least "serious"; a minor impairment is not sufficient.[200]

Ultimately, this leads to an assessment of the conflicting rights through a proportionality test. This assessment must weigh up whether less invasive means could achieve the desired success, taking into account mutual interests. It may be that a limitation of the data user group or pseudonymisation of the data is already sufficient.[201] This implies that if the processing of personal data continues for the specified purposes, the processing body must take appropriate technical and organisational measures, in particular

[197] *Paal*, in: Paal/Pauly Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 89 cip. 13 subseq.; *Caspar*, in Kühling/Buchner Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 89 cip. 4 subseq.
[198] see *Kamlah*, in: Plath, BDSG/DSGVO, 3. Aufl., 2018, Art. 17 cip. 19; *Herbst*, in: Kühling/Buchner Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 17 cip. 82
[199] *Kammann/Braun*, in: Ehmann/Selmayr, Datenschutz-Grundverordnung, Art. 17 cip. 63; also compare *Herbst*, in: Kühling/Buchner, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 17 cip. 82 who stresses that the impairment must in any case be at least "serious"
[200] *Herbst*, in: Kühling/Buchner, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 17 cip. 82
[201] *Dix*, in: Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht, marginal no. 35

to ensure the principle of data minimisation.[202] However, if storage for scientific purposes is no longer necessary, personal data must be deleted without undue delay (Art. 17 Sect. 1 lit. a GDPR).

III RISK ASSESSMENT

After having assessed the legal compliance of the data processing for research and statistical purposes in an urban traffic management environment, the third part of this DPIA consists of the actual risk assessment. The second and third data usage scenarios have only played a minor role in the previous legal compliance section, for example, to assess whether the collected data relates to data subjects or not, or regarding the potential impact of the processing on the data subjects' interests and their "reasonable expectations". In contrast, in this third part, all three usage scenarios play an equally important role. The reason for this is that the risk assessment in Art. 35 GDPR (also if it is just a scientific-legal risk assessment) does not have such a strong focus on the processing purpose as the compliance assessment, for example, with respect to the legal basis. Instead, it has a wider focus taking also unspecific risks into account (i.e. risks which are not implied by the processing purpose).[203] However, for the subsequent risk assessment, it is worth repeating once again that the risk assessment of this legal-scientific DPIA cannot go into detail, especially not on sources of risk, threats and countermeasures, since the target of evaluation cannot yet be sufficiently defined. However, this risk assessment helps to decide on which data protection risks should be focused when defining the appropriate data protection by design strategy. In the next sections, the specific risk assessment methodology will be first explained. The subsequent chapter gives an example for how the fundamental rights-based methodology can work in relation to the intended system. Finally, the legal-scientific DPIA promotes the implementation of the principle of responsibility and accountability principle as an overarching control of the risks that arise not only from the collection of the data, but especially from its later use.

1 Clarifying the risk assessment methodology

As mentioned in the introduction, the current DPIA refers to various risk assessment methods.[204] The following sections will clarify some ambiguities in relation to these methodologies.

1.1 Applying the processing principles in Article 5 GDPR to control data protection risks (esp. the likelihood of threats and severity of impacts)

As a first step, the risk assessment methodology proposed by the CNIL is very useful to answer the question of how to evaluate the likelihood and severity of data protection risks. In doing so, the CNIL differentiates between a threat causing a data protection risk and a data protection risk that can lead to a certain impact. An example of a threat may be when someone watches someone else's screen without their knowledge, for example on a train. Such a threat leads to the data protection risk in relation to the principle of confidentiality.[205] In turn, this data protection risk can have a negative impact on the individual concerned. For example, the person watching the other person's screen may see confidential information and use it against the other person in an employment context. The essential point is that the threat and impact can be assessed differently: On the one hand, one can assess a threat in terms of its likelihood.[206] In doing so, it is useful to examine the source of threat; for example, by differentiating between human and non-human factors, either from outside or within an entity; in regards of human

[202] *Nolte/Werkmeister*, in: Gola, Datenschutz-Grundverordnung, 2. Auflage 2018, cip. 47
[203] This is comparable to the (broad) approach applied to the definition of the scope, see above in section "II, 1. Application of data protection laws."
[204] See above in section "I. 2.3.1 Guidelines".
[205] See the example under https://www.cnil.fr/sites/default/files/typo/document/CNIL-PIA-2-Tools.pdf, pp. 18.
[206] See https://www.cnil.fr/sites/default/files/typo/document/CNIL-PIA-2-Tools.pdf, pp. 13–16.

factors, another distinction can be made as to whether the person is behaving in a random or intentional manner.[207] In either case, the question of the likelihood of threats can be answered on the basis of quantitative probability. On the other hand, one can focus on the impact with respect to its severity for the data subject,[208] and thus on the basis of a qualitative assessment. Indeed, one can – and must – additionally assess the likelihood that a threat leads to an impact.[209] For instance, the likelihood that a threat will lead to an impact of at least *one* data subject can be estimated by looking at the total number of data subjects (i.e. the more data subjects are concerned, the more likely it is that at least one data subject suffers an impact from the data processing).

In this framework, the implementation of the processing principles in Art. 5 GDPR (as well as all other requirements of the GDPR) by means of technical and organisational measures serves to reduce the corresponding data protection risks, more precisely the likelihood of threats (leading to an impact) and the severity of such an impact (if it may occur). In relation to data protection risks, two more aspects need to be clarified: First, this DPIA refers to all data protection risks that are consistent with the principles set out in Art. 5 GDPR. This appears to be an extended approach, compared, for example, to the CNIL risk assessment methodology. The reason for this is that the CNIL document "PIA -Tools" only lists the following events as risks: "Illegitimate access to personal data", "Unwanted modification of personal data", and "Disappearance of personal data".[210] Also the Art. 29 Data Protection Working Party only mentions these risks.[211] However, these risks primarily (but not exclusively) concern the data protection principles "confidentiality and integrity of personal data" (Art. 5 sect. 1 lit. f GDPR). Therefore, it seems that this approach does not take further risks with respect to the other principles into account. In contrast, the so-called Standard-Datenschutzmodell (in the following "SDM") of the Technology Working Group ("AK Technik") of the Conference of the German Data Protection Authorities ("Konferenz der unabhängigen Datenschutzbehörden des Bundes und der Länder") builds on "protection goals", which were developed by German data protection authorities and experts to operationalise the requirements better, which are set up by the data protection law.[212] Supported by the overarching goal of data minimisation, these data protection goals complement the already well-known IT security goals and create the following "fields of conflicts" between each other: intervenability against integrity, transparency against non-linkability, and availability against confidentiality.[213] On the one hand, regarding the explicit formulation, this approach apparently takes more data protection principles into account than the other approaches seem to do. On the other hand, these protection goals do not always appropriately cover all data protection principles. For example, the principle of accountability in Art. 5 sect. 2 GDPR is considered to be implemented through the protection goal of transparency. This is somewhat confusing because transparency is itself a processing principle under Art. 5 sect. 1 sent. 1 GDPR. However, if one goes into the details, the SDM actually covers all processing

[207] See https://www.cnil.fr/sites/default/files/typo/document/CNIL-PIA-2-Tools.pdf, pp. 10.
[208] See https://www.cnil.fr/sites/default/files/typo/document/CNIL-PIA-2-Tools.pdf, pp. 13–16.
[209] See *Bieker* et al., Die Risikobeurteilung nach der DSGVO, in: Datenschutz und Datensicherheit-DuD 42, 8 (2018), 492-496; as well as "Das Standard- Datenschutzmodell: Eine Methode zur Datenschutzberatung und -prüfung auf der Basis einheitlicher Gewährleistungsziele, Version 2.0b, von der 99. Konferenz der unabhängigen Datenschutzbehörden des Bundes und der Länder am 17. April 2020 (hereafter cited as SDM version 2.0b).
[210] See https://www.cnil.fr/sites/default/files/typo/document/CNIL-PIA-2-Tools.pdf, p. 22 (see regarding the synonymous wording "feared events", p. 12).
[211] See Article 29 Data Protection Working Party, Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679, 04 April 2017, 17/EN, WP 248, pp. 8, 13, as well as in the Annex 2 - Criteria for an acceptable DPIA., p. 21: "origin, nature, particularity and severity of the risks are appreciated (cf. recital 84) or, more specifically, for each risk (illegitimate access, undesired modification, and disappearance of data) from the perspective of the data subject".
[212] See *Friedewald* et al., Forum Privatheit, White Paper Datenschutz-Folgenabschätzung. 2017, p. 28. referring to Rost, Standardisierte Datenschutzmodellierung. In Datenschutz und Datensicherheit (DuD) 35, 6 (2012), p. 433-438.
[213] See SDM version 2.0b, p. 9-10; cf. *Friedewald* et al., Forum Privatheit, White Paper Datenschutz-Folgenabschätzung. 2017. 28. referring to Rost, Standardisierte Datenschutzmodellierung. In Datenschutz und Datensicherheit (DuD) 35, 6 (2012), p. 433-438.

principles listed under Art. 5 GDPR.[214] In fact, also the CNIL lists events that actually affect other data protection principles. For example, the feared event that the "data are used for purposes other than those planned and/or in an unfair manner (e.g. commercial purposes, identity theft, use against data subjects, etc.) or correlated with other information relating to data subjects (e.g. correlation of residence address and real-time geolocation data, etc.)" relates, primarily, to the principles of fairness and purpose limitation (and not – only or primarily – to the principle of confidentiality). Therefore, the current DPIA assumes that all of the approaches actually take into account all risks in relation to the data protection principles listed in Art. 5 GDPR. To avoid such ambiguities, the current DPIA will in any case directly refer to all data protection risks that are consistent with those in Art. 5 GDPR.

1.2 Categorizing the impact under the data subject's fundamental rights

In a second step, the CNIL and the German Technology Working Group assess, first and foremost, the severity of impacts by incorporating the distinction between physical, material and moral effects, as set out in recital 75 half sent. 1 GDPR. The CNIL, for example, distinguishes between four levels for each category, i.e. "negligible", "limited", "significant", and "maximum" impact and gives several practical examples for each category and level.[215] Highlighting the individual rights-based approach of data protection risk assessments, the German Technology Working Group considers that moral impacts are in particular unjustified violations of an individual's fundamental rights.[216] On this basis, the Working Group argues that the processing of personal data cannot only affect the fundamental rights to data protection and private life (Art. 7 and 8 ECFR) but also other rights such as the freedom of expression (Art. 11 ECFR) or the right to non-discrimination (Art. 21 ECFR). The intensity of a violation of these fundamental rights can then be categorised into three levels: light, medium and severe.[217] However, the fundamental right to data protection differs from the other rights as each processing of personal data automatically violates its scope.[218] This does not only mean that each processing of personal data must be justified, but also that the specific risk to that right is the non-application of its (data protection) principles.[219] It then depends on the specifics of the data and how the data is processed, whether the processing itself leads to a medium or severe violation of this right.[220] So far, it is clear that each processing of personal data falls within the scope of the protection of the right to data protection and that the processing may also infringe other fundamental rights, be it a minor, medium or serious infringement. However, the criteria (i.e. the nature of the data and how the data is processed) guiding the assessment of the intensity of a breach of the right to data protection seems less precise.

In order to further clarify the risk assessment, the current DPIA fits into the statement of Art. 1 sect. 2

---

[214] See the list in SDM version 2.0b engl, pp. 28 et seq., and explicitly at p.25 (p. 29 and p. 25 in the German version)
[215] See https://www.cnil.fr/sites/default/files/typo/document/CNIL-PIA-2-Tools.pdf, pp. 10-15.
[216] See SDM version 2.0b engl, p. 40 (p. 42 in the German version), referring to Kurzpapier Nr. 18 "Risiken für die Rechte und Freiheiten natürlicher personen" (see footnote 9 there), URL: https://www.datenschutzkonferenz-online.de/media/kp/dsk_kpnr_18.pdf, last access: 26 April 2020, cf. already *Friedewald* et al., Forum Privatheit, White Paper Datenschutz-Folgenabschätzung. 2017, p. 31.
[217] See Kurzpapier Nr. 18 "Risiken für die Rechte und Freiheiten natürlicher personen", URL: https://www.datenschutzkonferenz-online.de/media/kp/dsk_kpnr_18.pdf, last access: 26 April 2020; cf. *Friedewald* et al., Forum Privatheit, White Paper Datenschutz-Folgenabschätzung. 2017, pp. 32-33.
[218] See Kurzpapier Nr. 18 "Risiken für die Rechte und Freiheiten natürlicher personen", URL: https://www.datenschutzkonferenz-online.de/media/kp/dsk_kpnr_18.pdf, last access: 26 April 2020; cf. *Friedewald* et al., Forum Privatheit, White Paper Datenschutz-Folgenabschätzung. 2017, p. 31.
[219] See Kurzpapier Nr. 18 "Risiken für die Rechte und Freiheiten natürlicher personen", URL: https://www.datenschutzkonferenz-online.de/media/kp/dsk_kpnr_18.pdf, last access: 26 April 2020; cf. *Friedewald* et al., Forum Privatheit, White Paper Datenschutz-Folgenabschätzung. 2017, p. 33, referring to Bieker, F. (2018): Die Risikoanalyse nach dem neuen EU-Datenschutzrecht und dem Standard-Datenschutzmodell. In: Datenschutz und Datensicherheit (DuD) 42, Nr. 1.
[220] See Kurzpapier Nr. 18 "Risiken für die Rechte und Freiheiten natürlicher Personen", URL: https://www.datenschutzkonferenz-online.de/media/kp/dsk_kpnr_18.pdf, last access: 26 April 2020,); cf. *Friedewald* et al., Forum Privatheit, White Paper Datenschutz-Folgenabschätzung. 2017, p. 31.

GDPR that the regulation protects all the fundamental rights of the data subject. If one takes this statement seriously, the canon of all fundamental rights can serve as an objective scale for interpreting the GDPR, such as for the data protection risk assessment, in particular, for measuring the impact of data protection risks.[221] For doing so, the type of data, the context of collection and the purpose of the processing can serve as a very useful set of analytical instruments for the risk assessment. For example, referring to the type of data may help to define which substantial guarantee of the fundamental right to private life under Article 7 ECFR is specifically concerned. Referring to the context of the collection can help to assess whether the collection is actually covered by another right such as freedom (e.g. if the police collects personal data on the occasion of a political demonstration, this context may be covered by the freedom of assembly under Article 12 ECFR). And a close look at a new purpose may help assess whether subsequent processing causes, even if it takes place in the same context as before, a higher risk to such a right to freedom.[222] It is therefore less a question of assessing the risk "by reference to" these criteria,[223] [224] but rather of assessing the risk *against* these criteria.

Interestingly, the approach of taking the variety of all data subject's fundamental rights as an objective scale for assessing the impact of data protection risks does not conflict with the CNIL's methodology. The reason for this is that most of the examples listed in the CNIL's impact category can be attributed to one or another fundamental right, which are outlined in the following:[225]

- For instance, many of the examples for physical and moral impacts fall under the scope of the fundamental right to the integrity of the person (Art. 3 ECFR).
- The freedom to choose an occupation and to engage in work (Art. 15 ECFR) covers a missed career promotion or loss of employment.
- The receipt of unsolicited mails and targeted online advertising on a private aspect, which a data subject wanted to keep confidential, or even a damage of his or her reputation can violate, actually, the right to private life (Art. 7 ECFR).
- Financial losses and damages to property relates to the right to property (Art. 17 ECFR).
- A summons to court or criminal penalty can fall under the legal justice rights (Art. 47 et seq. ECFR).
- The freedom of movement and of residence (Art. 45) may cover a loss of housing or the inability to relocate.
- A lack of adequate care for a dependent person can fall under the rights of the child, the elderly, or handicapped persons (Art. 24–26 ECFR).

[221] See SDM version 2.0b, referring to Kurzpapier Nr. 18 "Risiken für die Rechte und Freiheiten natürlicher Personen" URL: https://www.datenschutzkonferenz-online.de/media/kp/dsk_kpnr_18.pdf, last access: 26 April 2020.
[222] See in more detail *von Grafenstein, M.*, pp. 493 et seq.
[223] See also SDM version 2.0b_engl, p. 41, where it still says that "[T]he likelihood and severity of the risk to the rights and freedoms of the data subject should be determined **by reference to** the nature, extent, circumstances and purposes of the processing; SDM version 2.0b, p. 42: "Eintrittswahrscheinlichkeit und Schwere des Risikos für die Rechte und Freiheiten der betroffenen Person sollten **in Bezug auf** die Art, den Umfang, die Umstände und die Zwecke der Verarbeitung bestimmt werden." [emphasis added].
[224] See SDM version 2.0a, p. 41 (p. 43 in the German version).
[225] Cf. the examples listed at https://www.cnil.fr/sites/default/files/typo/document/CNIL-PIA-2-Tools.pdf, pp. 13-16.

– Separation and divorce can be covered, for example, by the right to family life (Art. 33 ECFR).

Of course, it is open to discussion what exactly is covered by which fundamental right, if at all. In any case, these examples show that the diversity of all fundamental rights includes not only moral but also physical and material effects. More importantly, such an attribution of impacts to fundamental rights creates an important advantage compared to less fundamental rights-oriented methodologies: the reason for this is that this attribution makes it possible to systematise the variety of impacts according to an objective legal scale and to assess the severity of such an impact more reliably. Thus, the method provides more reliable results because the evaluation of the impact depends on the guarantee of the fundamental right in question and is therefore based on a normative rather than a quantitative-probabilistic assessment.[226] Moreover, this fundamental rights-based approach allows one to take the case-law into account that judicial courts have sometimes developed over decades, so that one can build on these results.[227]

Against this background, the Forum Privatheit rightly points out that such a *data protection* risk assessment must effectively reduce the risks against the fundamental rights of an individual.[228] Therefore, unlike a *normal* risk assessment, a data protection risk assessment cannot lead to the result that a risk for few data subjects is acceptable (whereas a risk for many data subjects is unacceptable). Such an approach would clearly clash with the idea of *individual* fundamental rights, meaning that *each* individual has fundamental rights that *must* be respected, irrespective of whether this individual is the only concerned data subject or whether there are more data subjects concerned by the processing. However, it is possible to prioritise and adapt protection measures according to the likelihood and severity of data protection risks to the fundamental rights of *an* individual.[229] How this can be done is illustrated in the following chapters.

1.3 Prioritising technical-organisational measures according to the risk

Focusing on the criteria catalogue proposed by the Art. 29 Data Protection Working Party,[230] this legal-scientific DPIA understands data protection risks as the element that links threats to a certain impact on data subject's fundamental rights. On this basis, the implementation of the processing principles of Article 5 GDPR by technical and organisational measures controls the data protection risks by reducing the likelihood of threats and the severity of impacts on one or more of the data subjects' fundamental rights. With respect to the intended system, the analysis above has shown that some data protection principles that are covered by Art. 5 GDPR are particularly relevant already to the legal compliance of the system. The data processing for research and statistical purposes in an urban traffic management environment can, in particular, only be based on the "legitimate interests"-clause of Art. 6 sect. 1 lit. f GDPR if the following data protection risks (for the following data protection principles) are met:

– The data must be pseudonymised to the greatest extent (in compliance with the **principle of data minimisation** under Art. 5 sect. 1 lit. c GDPR).

---

[226] See the criticism in *Friedewald* et al., Forum Privatheit, White Paper Datenschutz-Folgenabschätzung. 2017, p. 33; *Dijk* et al. A risk to a right? Beyond data protection risk assessments, in: Computer Law & Security Review 2016 32 (2), pp. 286–306.
[227] See *v. Grafenstein/Schulz*, The right to be forgotten in data protection law: a search for the concept of protection, Int. J. Public Law and Policy, Vol. 5, No. 3, 2015, pp. 247–269, p. 266.
[228] Kurzpapier Nr. 18 "Risiken für die Rechte und Freiheiten natürlicher personen", p. 6, URL: https://www.datenschutzkonferenz-online.de/media/kp/dsk_kpnr_18.pdf
[229] Cf. *Friedewald* et al., Forum Privatheit, White Paper Datenschutz-Folgenabschätzung. 2017, pp. 33-34.
[230] See Article 29 Data Protection Working Party, Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679, 04 April 2017, 17/EN, WP 248, Annex 2 - Criteria for an acceptable DPIA, p. 21.

– It must be ensured that the collected data cannot be combined with each other and with data from outside the system in an unlimited, i.e. uncontrolled, way (in compliance with the **principle of purpose and storage limitation** according to Art. 5 sect. 1 lit. b and e GDPR).

– The data processing must be as transparent as possible for the data subjects concerned (concurrent to the **principle of transparency** under Art. 5 sect. 1 lit. a GDPR).

It is possible to reduce these data protection risks by implementing the corresponding data processing principles through appropriate countermeasures (Art. 35 sect. 7 lit. d GDPR). The SDM, for example, lists in respect of the protection goals "data minimisation", "non-linkability" (corresponding to the purpose limitation principle) and "transparency", the following technical and organisational measures:

Protection Goal	Examples for measures
Data minimisation	- Reduction of collected attributes of the data subjects - Limitation of access to personal data - Preference for *automated* data processing (not decision-making!) - Pseudonymisation and anonymisation processes - Definition and application of deletion concepts - Rules regarding the control of processing changes
Purpose limitation	- Limitation of processing, usage and transfer rights - Closure of application programming interfaces - Separation according to organisational structures - Separation according to roles-based rights attribution - Purpose-oriented pseudonymisation and anonymisation - Defined processes for purpose changes
Transparency	- Information, specification, documentation of data - Specification, documentation of system and activities - Information, specification, documentation of processes

The examples show that there is a considerable overlap of measures that implement both the principle of data minimisation and purpose limitation. The same is actually true with respect to many other protection measures: for instance, D'Acquisto illustrated how pseudonymisation can help implement all processing principles listed under Article 5 GDPR.[231] Such an overlap means that controllers have to prioritise and adapt the protection measures according to the likelihood and severity of the data protection risks for the fundamental rights concerned.[232]

2 In the intended system, counterfeiting...

In applying this fundamental rights-based approach, the implementation of the data protection principles thus *mediates* the impact of a data processing on (potentially, all) the fundamental rights of the data subject. The next sections demonstrate this approach examining the potential impact of the intended system on specific fundamental rights. This assessment will be brief, as most aspects have been

[231] See G D'Acquisto's presentation at the workshop on "Pseudonymization as a data protection by design instrument" at https://www.enisa.europa.eu/events/uld-enisa-workshop/uld-enisa-workshop-pseudonymization-and-relevant-security-technologies.
[232] Cf. *Friedewald* et al., Forum Privatheit, White Paper Datenschutz-Folgenabschätzung. 2017, pp. 33-34.

discussed previously.

2.1 … insights into private life

The processing of the personal data for research and statistical purposes in an urban traffic management environment affects the fundamental right to private life of all data subjects concerned.[233] The severity of this impact depends on the extent and depth of insights gained through the data processing. The broader and deeper these insights become, the more severe the impact is on the individual's fundamental right to private life. The likelihood of the underlying threat – which in this context means in particular (but not exclusively) the creation of movement patterns of the data subjects – is considered very high, as the processing purpose necessarily implies the creation of these movement patterns. In particular, the implementation of the principle of data minimisation can reduce the severity of such an impact on the individual's' right to privacy.[234] Appropriate measures may be the anonymisation and pseudonymisation of the collected data. Of course, the implementation of this measure must be balanced against the interests of the controllers and the public in the data processing.[235]

2.2 … legal enforcement by private and public bodies

The intended data processing can also impact the data subjects' right to private autonomy and his/her right to a fair trial. For example, when a public order agency processes the collected data by the parking lot sensors by using it against an individual in relation to a traffic road law offense, this could conflict with the individual's judicial right to a fair trial.[236] In contrast, when a private insurance company uses vast amounts of collected data to an individual, for example, by unfairly increasing the fees that this individual has to pay for a traffic accident insurance, this could conflict with the individual's right to private autonomy.[237] The third data usage scenario demonstrates that such an impact is possible.

Before such a risk becomes specific, i.e. the data is actually used in such a manner, it is difficult to determine the severity of such an impact. Whether a data subject might lose their insurance as a whole or might just pay a higher price, or whether he or she is sentenced to jail or just to pay a fine, depends on the specific circumstances of the later case. It is therefore necessary to check at this later stage whether and, if so, under what conditions the data will actually be used for such a purpose. In contrast, it is easier to assess already in advance the likelihood that a particular threat will affect an individual's judicial rights. One factor can be in this regard the amount of collected information. The idea behind this is that the more information is collected, the more likely it will be that at least part of the information will be used against an individual one day. Another factor may be the motivation of individuals and/or entities to use the collected data in the course of a legal proceeding against an individual. Again, with respect to the target of evaluation of this DPIA, the data shall not be processed for that purpose but only for research and statistical purposes in an urban traffic management

---

[233] See Art. 7 ECFR.

[234] This seems to be the reason why the European Court of Justice always refers to the right to private life instead of the right to data protection, when it requires that the data processing must be strictly limited to what is absolutely necessary, and why it does not matter whether the data is sensitive or the individual was inconvenienced in any way. See ECJ C-293/12 and C-594/12 (Digital Rights vs Ireland), cip. 33 and cip. 52; affirmed in the subsequent case of "Schrems vs. Ireland", ECJ C-362/14, cip. 92. This also explains why the Court applies this requirement for both the data collection by public and private entities –– see ECJ C-293/12 and C-594/12 (Digital Rights vs. Ireland), cip. 52; ECJ C-473/12 (IPI vs. Englebert), cip. 39; ECJ C-92/09 and C-93/09 (Schecke vs. Germany), cip. 77 und 8; ECJ C-73/07 (Satakunnan Markkinapörssi und Satamedia), cip. 56 –– the requirement does hence not result from the principle of proportionality of law, at least, not exclusively because this is directly applicable to public sector.

[235] See above in section "II, 2.2.2. 'Legitimate interests'-clause (Art. 6 sect. 1 lit. f GDPR)".

[236] See, for example, Art. 47 et seq.

[237] See, for example, the case, ECJ C-250/97 (which was, however, decided before the European Charter has come into force).

environment. However, the third data usage scenario illustrates that the motivation of public authorities (e.g. a public order agency or the police) and private bodies (e.g. insurance agencies) can be very high. In this respect, this DPIA considers the likelihood that the collected data will someday be used for such purposes between medium and high.

To reduce the severity of a potential impact as well as the likelihood of threats, it is necessary to implement data protection principles of data minimisation and purpose limitation, for example, through pseudonymisation, functional data separation and mechanisms of access and usage control. These measures can significantly reduce the data protection risk that instead of only using it for research and statistical purposes, the collected data is used for law enforcement purposes. If the data should one day be used for such law enforcement purposes – provided that such a usage would be lawful – further data protection principles apply. In particular, the principle of fairness (and of "intervenability", which may be one sub-goal of the fairness principle) reduces the risk that the data is used unfairly against a person. The principle of accuracy also becomes more important. The reason for this is that an individual's judicial rights require that the information used against an individual is at least correct.[238] In particular, these principles could be met by properly implementing the individual's rights to correct and delete (or restrict) data (processing) and by establishing a single point of contact for the data subject.[239]

2.3 … the feeling of being "under constant surveillance"

The intended data processing for research and statistical purposes also impacts the substantial guarantee that data subjects have the right to know, at least to some extent, what other people know about them. As mentioned previously, in the case of "*Digital Rights vs. Ireland*", the European Court of Justice stated that the collection, storage and subsequent use of personal data without informing the data subjects "is likely to generate in the minds of the persons concerned the feeling that their private lives are the subject of constant surveillance."[240] Whether such a guarantee is protected by the right to data protection under Art. 8 ECFR, the right to private life under Art. 7 ECFR or the right to the integrity of the person under Art. 3 sect. 1 ECFR of the individual concerned is open for discussion. In any case, the severity of such an impact depends on the extent and degree of opacity of the data processing. The likelihood of threats that can lead to such an impact is considered very high. On the one hand, the research and statistical purposes in an urban traffic management environment can imply the largest possible amount of processed data, depending on the more specific research question. On the other hand, the current DPIA has come to the conclusion that the controllers have to inform the data subjects in accordance to Art. 14 GDPR, instead of to Art. 13 GDPR. This means that the controllers must additionally inform the data subjects about the data categories and sources, but the controllers also profit from the legal privileges when and how they have to give the information. In particular, the controllers do not have to inform the person immediately at the time of collection but may later pass the information to the individual concerned or pass it onto the public. This increases the likelihood that processing will remain significantly opaque.

The implementation of the principle of transparency therefore can — following the terminology of the German Constitutional Court — pursue the specific aim "to diminish the threat resulting from the lack of knowledge about the real relevance of the data, to counter unsettling speculations, and to enable the data subjects concerned to question these measures in a public discourse."[241] Although the considerations of the German Constitutional Court cannot be used, of course, for the *interpretation* of European law,

[238] See *Frenzel* in Paal/Pauly Art. 5 cip. 39.
[239] See the examples in *Friedewald* et al., Forum Privatheit, White Paper Datenschutz-Folgenabschätzung. 2017 p. 36.
[240] See ECJ C-293/12 and C-594/12 (Digital Rights vs Ireland") cip. 37.
[241] See BVerfG, 2nd March 2010, 1 BvR 256/08, 1 BvR 263/08, and 1 BvR 586/08 (Data Retention), cip. 241.

they can serve as a source of *inspiration* for the question of the implementation of the principle of transparency. To counterfeit such an unspecified fear (or *Angst*), controllers must also apply the principles of data minimisation and purpose limitation, and the corresponding measures as previously described.

2.4 … discriminatory effects

Last but not least, the data processing of the intended system can also impact the rights to non-discrimination of the data subjects. Again, based on the likelihood of such a threat, the processing for research and statistical purposes in an urban traffic management environment does not necessarily mean that these data subjects are in fact treated differently. However, the second usage scenario has at least shown that data processing may potentially lead to such an impact. This may be the case in particular when people using public and/or private means of transportation are treated differently, depending on how much they pay for it or the data they want to disclose. The current DPIA considers such a likelihood to be between medium to high. The reason for this increased likelihood is that even if the research and statistical purposes do not necessarily imply such discriminatory effects, these effects are potentially inherent in the system. On the contrary, the DPIA regards the severity of such an impact only as low to medium, as long as the research and statistical activities are not based on criteria listed in Art. 21 ECFR.[242] Nevertheless, this DPIA sees an impact on these rights to non-discrimination already when people are just treated differently, even though it is not yet clear whether the discrimination is justified or not. The reason for this is that an answer to the question of whether a discrimination is justified or not cannot be often determined from an objective point of view but is the result of a socio-political negotiation process.[243] To enable such a process, the criteria used to distinguish data subjects from each other must be made transparent. This makes it possible to find out whether the data processing ultimately discriminates or not. Thus, in this context, it is less the implementation of the principle of fairness but transparency that controls the impact of unjustified discriminations that are caused in the second usage scenario.

2.5 … by applying the principles of data minimisation, purpose limitation and transparency (in particular)

The preceding assessment has not yet decided on the final risk. Such a final decision would require the following: Firstly, to decide on how likely it is that a certain threat impacts each substantial guarantee and how severe this impact is. Secondly, to implement the data protection principles through the necessary measures in order to reduce the risks, mediating between threats and impacts. Lastly, coming up with a conclusion about the risk that remains after all. This DPIA, on the other hand, did not decide on the final impact on the respective fundamental rights nor on the necessary countermeasures. Instead, this DPIA provides for a structure that can serve as a basis for an appropriate data protection by design strategy. To achieve this task, the DPIA has come to the result that, among all the data protection principles listed under Art. 5 GDPR, the data protection principles of data minimisation, purpose limitation and transparency are in particular important for controlling the data protection risks caused by the proposed system. All three principles can particularly help to protect the data subject in exercising their following fundamental rights: to private life, internal freedom of development (i.e. against the feeling of being under "constant surveillance"), private autonomy and a fair trial and non-discrimination.

[242] If the data processing builds on criteria listed under Art. 21 ECFR, criteria Art. 9 GDPR provides for a special regulation.
[243] See the discussion on the panel "Exploring the 'Design' in Privacy by Design" at CPDP2018, taking the example of a "Transparent Charging Station" given by Harm van Beek, online available at https://www.youtube.com/watch?v=qMwQRB7nJwo.

The principle of data minimisation can help to effectively protect the data subjects' right to private life, in particular by making data anonymous and, where this is not possible or not feasible for the specific research purposes, by pseudonymising the collected data as far as possible. This will prevent the collected data from revealing personal aspects of the private life of the persons concerned (which is protected not only in the home of the person concerned but also in public). Similarly, the principle of data minimisation protects the data subjects' rights to private autonomy and to a fair trial. The less data is collected that relates to a data subject or the more difficult it is to link that data to them, the less a controller or another party can use this type of data against the data subject, whether in relation to the conclusion or performance of a contract or another legal proceeding. Consequently, the principle of data minimisation also protects the data subjects against the feeling of being under "constant surveillance", as surveillance is limited to the extent in that the collected data does not relate to an individual (or makes it at least more difficult to relate to it).

The same idea applies to the purpose limitation principle, which in principle, prevents the data controller(s) from processing the data for a purpose other than originally intended. This principle protects the right to private life of the data subject, in particular if this principle is interpreted in a way so that the subsequent processing may not disclose more aspects of the private life of a data subject than initially stated.[244] Similarly, the principle of purpose limitation can protect the data subject's right to private autonomy or a fair trial. If the purpose of the data collection did not specify the subsequent use of the collected data, for example, in relation to the conclusion or performance of a contract or a legal proceeding, the purpose compatibility test provides for an effective mechanism that controls such a purpose change (e.g. prohibiting the purpose change or allowing it only under certain conditions). As a consequence of such restrictions on subsequent processing, the principle of purpose limitation also helps protect the data subject against the feeling of being under "constant surveillance".

Last but not least, the principle of transparency can contribute to the data subjects' right to private life because effective information about the intended data collection enables data subjects to prepare and respond in accordance to their preferences. For example, some data subjects may wish to have the MAC address(es) of their device(s) automatically filtered out at the time of collection. This is only possible if these data subjects are properly informed. The same idea applies to the data subject's rights to private autonomy and a fair trial. If a data subject is informed early enough that the data should be used against them, for example, in relation to the conclusion or performance of a contract or a legal proceeding, the data subject is able to respond accordingly (e.g. by correcting the processed data or rejecting the use altogether). Finally, the principle of transparency can also protect the right to non-discrimination of the data subjects. Even if the collected data is not processed in accordance to the criteria listed under Art. 21 sect. 1 EuChFR (sex, race, colour, and so on), the processing may conflict with this right if the data processing leads to an arbitrary discrimination (i.e. if the discrimination cannot be justified on reasonable grounds). In fact, whether the reason for discrimination is arbitrary or not depends on values that very often still need to be defined in the course of socio-political negotiation processes. In this regard, the principle of transparency can make the inherent logic of the data processing transparent enough so that the affected data subjects can — beside (or even instead) of their representatives and the legislator — discuss whether they agree with the specific logic or not.

[244] See *v. Grafenstein, M.*, The Principle of Purpose Limitation: The Risk-Based Approach, Legal Principles and Private Standards as Elements for Regulating Innovation, 1st ed. (2018), Mohr Siebeck, pp. 87 subseq.

3 Responsibility and accountability as an overarching risk control (focusing on how data is used, besides its collection)

3.1 Challenges of contextual usage control

In any case, a major challenge for implementing the principle of data minimisation and purpose limitation as well as adjusting the necessary information is that the effectiveness of the protection depends on how these principles are applied when the data is used, sometimes long after it has been collected. The following paragraphs shall illustrate this challenge by ultimately focusing on the implementation of the principle of responsibility and accountability as a risk control supporting the other processing principles.[245]

3.1.1 How to guarantee pseudonymisation?

With respect to an effective anonymisation or pseudonymisation of personal data, the UK Anonymisation Network (UKAN) emphasised in its study *The Anonymisation Decision Making Framework* that "we have to deploy effective anonymisation techniques and assess re-identification risk in context".[246] In this regard, the Art. 29 Data Protection Working Party refers to the so-called "content" element. It should be recalled that this criterion does not require that the data itself is sufficient to identify a person. This criterion rather includes situations where additional information is necessary.[247] In fact, the Working Group also refers to other elements, in particular to the so-called "result" element, which means that data can relate to an individual even if it is not about them but because they affect the individual. At first sight, these situations therefore do not seem to fall within the scope of anonymisation (or pseudonymisation) techniques. At second glance, however, the UKAN also refers to such situations. The UKAN calls the respective category "secondary personal data". This is data in which the data units are not people but facts that are likely to be related to a certain person. The UKAN cites the example of a fire that may occur somewhere and, if it does occur, is registered by the local authorities. This data is not about people but *can* refer to people: While the fact that a fire took place somewhere in the forest is not personal, the fact that a fire took place in someone's home is personal. In fact, the UKAN itself admits that the decision as to whether or not such object-related information relates to people "can seem a little tricky."[248] An important criterion for making such a decision are the consequences that an individual may suffer because of the use of that information.[249] It is this precisely this last criterion which is the reason why in this DPIA the category "secondary personal is considered to correspond to the "result" element used by the Art. 29 Data Protection Working Party. Whether data is "secondary personal" depends on whether the data usage "is likely to have an *impact* on a certain person's rights and interests, taking into account all the circumstances surrounding the precise case [emphasis added]."[250]

This DPIA has come to the conclusion that all types of collected data must be considered personal, even if it has been "anonymised" in a first step. The reason for this strict and therefore, rather broad approach is that all data can always be related to an individual in a second step by adding further information. Firstly, this is the case with respect to the movement patterns created on the basis of the collected Wi-Fi

[245] Cf. also SDM version 2.0b, p. 13.
[246] See *Elliott* et al. "The Anonymisation Decision Making Framework", Ukan publications, 2016, p. vi.
[247] See DPIA, p. 25, referring Article. 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007, 01248/07/EN, WP 136, pp. 10 and 13.
[248] See *Elliott* et al., "The Anonymisation Decision Making Framework", Ukan publications, 2016 p. 9, footnote 18.
[249] See *Elliott* et al., "The Anonymisation Decision Making Framework", Ukan publications, 2016, p. 9, footnote 18.
[250] See Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007, 01248/07/EN, WP 136, p. 11.

data. In this regard, even if the collected MAC addresses were at least pseudonymised, these movement patterns always serve as identifiers. The combination with further information can lead to situations where all information associated with this pattern relates to a specific individual (e.g. when someone recognizes the movement pattern of his wife commuting to and from her daily work). Secondly, even data collected by parking sensors that do not appear to be personal actually relate to an individual if appropriate additional information is added. For example, the moment a public authority compares information about the parking location of certain vehicles with its information about "parking prohibited" areas, such an authority can link that information to a car owner at the point where road traffic law has been violated. Thirdly, the same actually applies with respect to the CCTV camera data that seems to be completely "anonymised" since nobody can recognise the recorded persons and/or cars. However, as long as the recorded data still show a certain action taken at a particular location and time, that information can later be related to a person. An example of where this might be the case is a witness who cannot say what exactly happened at a certain time and place but can testify that a certain person was there (and only this person). Both parts of information about the action and the person can then be combined, which means that it must have been that person who acted in the recorded manner.[251] If these types of data relate to an individual because of the "content" element or only because of the "result" (or also "purpose") element, these examples demonstrate that all collected data by the intended system relates in one way or another to an individual. This leads us to conclude that the only way to prevent, for example, a witness from (perhaps erroneously) recording information about a particular action relating to a person is to prohibit the collection of that information or at least to require the recording to be deleted immediately.[252] When collecting such information (and not deleting it), data protection laws apply. However, the DPIA has come to the conclusion that even if data protection laws apply, one has to take the different levels of risk into account that the data is "de-anonymised".[253]

This approach corresponds to the discourse about the trade-off between anonymisation and the usefulness of data. In this debate, some people argue that, in fact, there is no data that is both usable and completely anonymised (i.e. without the remaining risk that the data could be de-anonymised). It is therefore always necessary to weigh the benefits of data against the risk that the data is de-anonymised.[254] The first step in this balancing exercise is to measure the de-anonymisation risk. Fortunately, in the last decades, the debate was able to elaborate on methods that statistically calculate the probability (i.e. the risk) that anonymous data is de-anonymised. In any case, even though such progress has been made, the UKAN criticises that so far the discussion has mainly focused on the "statistical properties of the data to be released/shared" but has not adequately considered the context in which the data is released or shared.[255] Building on these existing approaches of "statistical disclosure control", the UKAN advocates adding contextual factors. In particular, these factors may be:[256]

- How de-anonymisation may occur accidentally, i.e. without malicious intent
- The motivation of potential "attackers" who deliberately seek to de-anonymise the data
- Information that originate outside of the original data set and may be linked with the original

[251] See DPIA, pp. 34.
[252] Cf. DPIA, pp. 34, referring to the Article 29 Data protection Working Party, Opinion 5/2014 on Anonymisation Techniques, adopted on 10 April 2014, 0829 /14/EN WP 216, pp. 11 and 12.
[253] See DPIA, pp. 34.
[254] See *Elliott* et al., "The Anonymisation Decision Making Framework", Ukan publications, 2016 pp. 18-21, referring to Ohm, P. (2010) Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization; UCLA Law Review, 57: 1701-1777, available at: http://www.uclalawreview.org/pdf/57-6-3.pdf
[255] See *Elliott* et al., "The Anonymisation Decision Making Framework", Ukan publications, 2016, p. 21.
[256] See *Elliott* et al., "The Anonymisation Decision Making Framework", Ukan publications, 2016, pp. 21 and 22.

data

– The governance structures for managing the access and usage of the anonymised data that may influence the de-anonymisation risk

This approach, called "functional anonymisation", combines solutions that focus on the data properties as such, with other controls that take the environmental factors of the data usage into account. The UKAN lists as data-centric solutions, for example: at the data meta-level, sampling, change of variables and decrease of the level of detail; and as solutions that distort the data itself: data swapping, overimputation, rounding, cell and value suppression, and k-anonymisation.[257]

Looking at the environmental-focused solutions, the UKAN elaborates on the control categories on the basis of the following three questions: First, who has access to the data in question? Second, what kind of analysis is allowed? Third, where and how (i.e. under what conditions) can the data be processed?[258] The first control category follows the logic that there is a connection between the level of risk and the type of entity using the data as well as its role in society. For example, there may be a lower risk if the user is associated with an institution that has already implemented certain risk controlling infrastructures than if the user is acting on a stand-alone basis. Correspondingly, the risk may be reduced even if the user did not belong to an established institution if he or she had undergone a specific training process to control the risks themselves. The second question aims to control the data processing per se, for example, by banning certain types of analysis or requesting a project approval process. In this second regard, the UKAN refers to the UK Administrative Data Research Network, which includes a formal project approval panel. The control of certain processing outputs also falls under this category. This could affect the type of information collected.[259] However, it is also possible to control the purposes for which the collected information can be used. This could be done, for example, by providing a whitelist for processing purposes that are unproblematic, a blacklist for problematic purposes and unrecorded purposes for which the controller, as previously mentioned, must undergo a formal approval process.[260] The variety of processing controls leads us to the third category regarding the "where" and "how" of the data processing. The UKAN distinguishes between four modes: Open access, i.e. the data is made available to the general public; delivered access, where the data is delivered to a specific user; on-site security settings, which are considered to be the most secure but also the most restrictive mode (because the user is able to use the data only within the premises and with the facilities provided by the controller); finally, virtual access, currently being discussed as "the future of research access" (because it combines the security level of on-site settings with more flexibility for the user who can to a certain extent use its own facilities).[261]

In all four modes, the controller can grant access to the data by more or less requiring the user to meet certain conditions. These conditions may apply to all three control categories, i.e. who specifically gets access to the data for what kind of processing and under what conditions. Of course, such conditions can be offered at different levels depending on the risk. If certain data sets are at high risk, the controller may grant access to those data sets under more stringent conditions than at lower risk.[262] Thus, the access scheme depends on a de-anonymisation risk assessment. In this assessment, the UKAN recommends taking the following three factors into account: the sensitivity of the data, the data

[257] See *Elliott* et al., "The Anonymisation Decision Making Framework", Ukan publications, 2016, pp. 43-52.
[258] See, *Elliott* et al., "The Anonymisation Decision Making Framework", Ukan publications, 2016, pp. 52 and 53, referring to Duncan, G.T., Elliot, M. J., & Salazar-González, J. J., "Statistical Confidentiality." New York: Springer, 2011.
[259] See *Elliott* et al., "The Anonymisation Decision Making Framework", Ukan publications, 2016, p. 54.
[260] Cf. *Elliott* et al., "The Anonymisation Decision Making Framework", Ukan publications, 2016, p. 55.
[261] Cf. *Elliott* et al., "The Anonymisation Decision Making Framework", Ukan publications, 2016, pp. 56-58.
[262] Cf. *Elliott* et al., "The Anonymisation Decision Making Framework", Ukan publications, 2016, p. 58.

disclosiveness and the environmental risk. For example, if the data is very sensitive, one must re-balance this higher risk for the data subject by providing stricter disclosure and environmental controls. In contrast, environmental controls may be lower if the disclosure risk is low and the data is not sensitive and so on. The entire de-anonymisation risk for the data subject must be finally weighed against the interest of the controller(s) in disclosing and/or processing the data.[263]

3.1.2 How to guarantee purpose limitation?

Similar to the approach of "functional anonymisation", the principle of purpose limitation also requires the data controller(s) to consider the context of the subsequent usage of the data. There are two reasons for this: The first reason refers to the first component of the principle of purpose limitation, which obliges the controller to specify the purpose of the subsequent processing at the latest at the time of collection of the data. The purpose therefore to a future situation in which the processing is to take place.[264] This reference (to a future situation) is necessary because the relevance of data cannot only be determined by the data properties itself, but must be assessed with respect to the context in which the data is used.[265] Given that the relevance of data depends on its usage context, the specification of the purpose is a necessary precondition for assessing whether and, under which conditions a certain processing activity can be legally justified.[266] This is one reason why the usage context is particularly relevant for implementing the principle of purpose limitation. The second reason refers to the second component of the principle of purpose limitation. As already explained in the DPIA, this second component allows the data controller to process personal data for a purpose other than that for which the data was collected, and only if it is not incompatible with its original purpose. This so-called purpose compatibility assessment, therefore, focuses on the usage context and compares this situation with the data situation that has formerly been indicated by the original purpose.[267]

The DPIA does indeed focus on the necessity and proportionality of the intended data processing for research and statistical purposes in an urban traffic management environment. In this regard, the principle of purpose limitation does not pose a problem for a data controller as long as the controller only processes the collected data for the research and statistical purposes in an urban traffic management environment. If this data has already been collected for these research purposes, there is no change of purpose, which would require a purpose compatibility assessment. In principle, there was only a problem if the controller processed personal data not collected for these purposes, in particular, personal data originating from outside the system (e.g. from social networks). Such a change of purpose could indeed conflict with the principle of purpose limitation. However, for research and statistical purposes, Art. 5 sect. 1 lit. b) half sent. 2 GDPR provides a special legal privilege, which states that the data processing is not incompatible with the original purposes as long as the controller(s) process(es) this data only for research purposes and the safeguards under Art. 89 sect. 1 GDPR are met. Thus, the main question is how to guarantee that personal data – collected within the intended system or outside of it – are processed only for these research and statistical purposes. The SDM lists in version 2.0b several technical and organisational measures that can help to implement the principle of purpose limitation, such as pseudo- and anonymisation, attribute-based credentials, functional separation of data and

[263] Cf. *Elliott* et al., "The Anonymisation Decision Making Framework", Ukan publications, 2016, p. 42.
[264] See *v. Grafenstein, M.*, The Principle of Purpose Limitation: The Risk-Based Approach, Legal Principles and Private Standards as Elements for Regulating Innovation, 1st ed. (2018), Mohr Siebeck, 2017, pp. 420 subseq.
[265] Cf. BVerfG, 15th of December 1983, 1 BvR 209, 269, 362, 420, 440, 484/83 (Volkszählungsurteil), cip. 158 and 159.
[266] See Article 29 Data Protection Working Group, Opinion on purpose limitation, 00569/13/EN WP 203, adopted on 2 April 2013, pp. 51 seq.
[267] See *v. Grafenstein, M.*, The Principle of Purpose Limitation: The Risk-Based Approach, Legal Principles and Private Standards as Elements for Regulating Innovation, 1st ed. (2018), Mohr Siebeck, pp. 485 subseq.

systems, identity management, audits and roles-based rights attribution.[268] The examples show that there is a considerable overlap of measures that implement both the principle of data minimisation and purpose limitation. Among these measures, the functional separation of data and systems should be discussed in more detail. The Art. 29 Data Protection Working Party highlights the relevance of the concept of functional separation of data, especially in situations where multiple parties are involved, and the data cannot be fully anonymised.[269]

The concept of functional separation of data requires the involved parties to organise the system in a way to guarantee that each party can only process the data that is absolutely necessary for its purpose. This means with respect to the intended system that the concept of functional separation of data may be implemented as follows: KiwiSecurity processes the parking lot data only to provide the parking lot service, which means that KiwiSecurity is, in this regard, the controller. Cisco processes the Wi-Fi data only to provide the Wi-Fi service and is the controller in that respect. Thus, the system must be technically and organisationally designed in such a way that these types of data are only processed for these purposes and are not combined with each other. If these data are to be additionally processed in an urban traffic management environment for research and statistical purposes, technical and organisational measures must guarantee that only the party who determines the research purposes has access to the data only in order to process the data for these purposes. So, if Cleverciti does not collect the CCTV camera data for its own purpose(s) but only for the research purposes being specified by another entity, Cleverciti is not the controller but the processor acting on behalf of this other entity. In this respect, Cleverciti must "not process those data except on instructions from the controller" (Art. 29 GDPR) and "implement appropriate technical and organisational measures to ensure a level of security", in particular, "ongoing confidentiality, integrity, availability and resilience" of the processing systems and services (Art. 32 GDPR). These requirements, therefore, aim to guarantee – again at a technical and organisational level– that the data cannot be (mis)used for a different purpose and/or another entity than specified by the controller. This approach applies any kind of new processing purpose to any other entity involved in the intended system and thus to any person working for such an entity and coming into contact with that data.

3.1.3 How can further measures be adapted to the constant changes?

The security by design-requirement under Art. 32 GDPR – which applies to both the processor and controller – leads us to further legal obligations, in which the context of the subsequent use of the collected data plays a crucial role. As mentioned previously, the data controller and processor have to guarantee the *ongoing* confidentiality and integrity of the collected data. Similarly, the data protection by design-requirement under Art. 25 GDPR requires the controller – in this regard the processor is not the regulation addressee of this provision – to implement the appropriate measures "both at the time of the determination of the means for processing and at the time of the processing itself". So, whether the measures are appropriate depends not only on an assessment at the time of data collection but also on the subsequent use of the data.[270]

[268] See SDM version 2.0b, p. 33 (p. 33-34 in the German version); cf. already *Friedewald* et al., Forum Privatheit, White Paper Datenschutz-Folgenabschätzung. 2017, p. 34.
[269] See Article 29 Data Protection Working Group, Opinion on purpose limitation, 00569/13/EN WP 203, adopted on 2 April 2013, p. 19.
[270] See *Martini*, in: Paal/Pauly, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 32 cip. 55; Art. 24 cip. 31 subseq.; *Hladjk*, in: Ehmann/Selmayr, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 32 cip. 2.

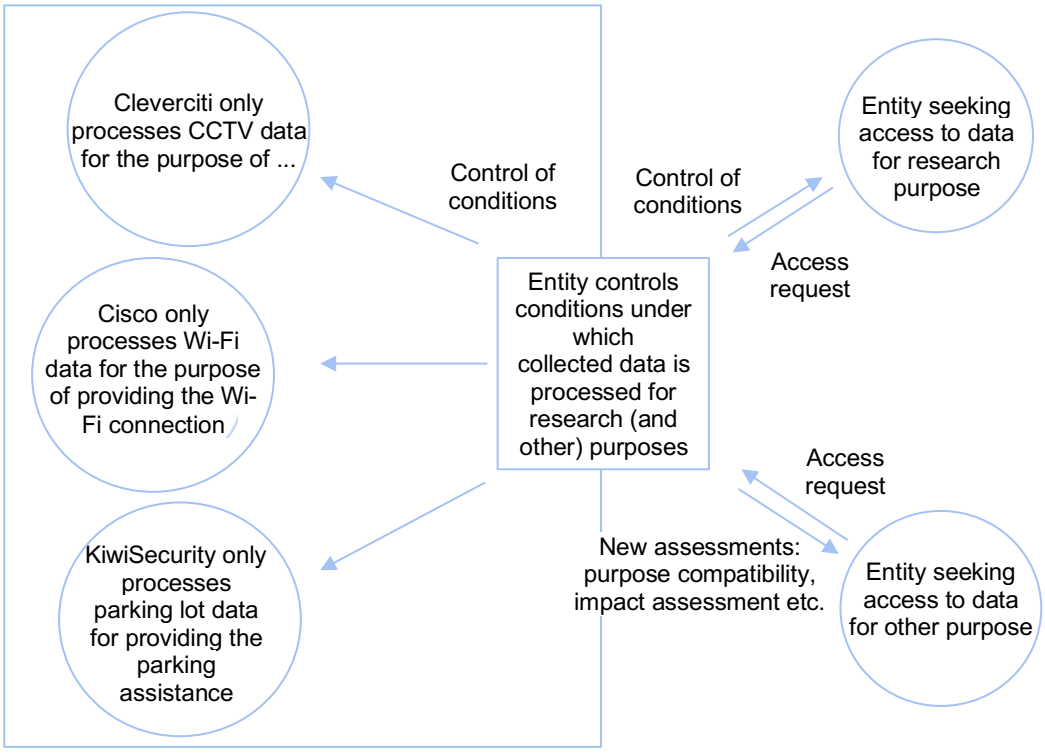
3.2 Codes of conduct and certification as control mechanisms

One way to guarantee that the data processing principles are adequately met, especially at a later stage, is to set up a code of conduct and/or certification scheme.[271] The reason for this is that codes of conduct and certification schemes are, in light of their institutionalisation, particularly suitable to control the intended data processing not only at the moment the data is collected but also at a later stage. After a more detailed explanation of this function, the next paragraphs will briefly summarise the main advantages of codes of conduct and certification scheme as well as their main differences.

3.2.1 Common functionalities

Codes of conduct and certification schemes are both co-regulatory instruments designed to specify the provisions of the GDPR and to control its application.[272] While the specification is provided by private parties (e.g. data controllers and/or processors) in conjunction with the competent Data Protection Authority (DPA),[273] specific bodies that have to be accredited by the competent DPA, are primarily responsible for monitoring compliance with the code of conduct or certification scheme.[274] Such an institutionalised control mechanism makes codes of conduct and certification schemes particularly useful for controlling how the collected data is used. The reason for this is that data subjects generally do not have the necessary capacities and practical means to check that the controller(s) and processor(s) apply the GDPR with respect to the data concerning them.[275] In particular, data subjects usually lose their actual ability to control "their" personal data after this data has been collected. This is all the more the case when data is only considered personal because someone else may relate it later to an individual. The following example will illustrate this case: The DPIA assessed the case in which CCTV cameras recorded a certain human action at a particular time and place, but this information was pseudonymised (irreversibly) in such a way that no one could see (anymore) who that person is. However, it is possible that somebody else (let us say a "witness") claims to have seen this particular person at this time and place but has not seen what this person did. The information about the recorded action is therefore attributed to the individual. If this attribution is incorrect (for example, because the witness mistakenly remembers the situation), the affected person was certainly unable to control the data situation at the time of collection. In this case, the data subject has a particularly strong need for this type of use to be controlled by another entity. The following picture may illustrate this idea in view of the above considerations regarding the functional separation of personal data:

[271] In contrast, binding corporate rules do not play a role in this study because their legal effects mainly concern the transfer of personal data to third countries, see Art. 46 sect. 2 lit. b) in combination with Art. 47 GDPR.
[272] See recital 89 and 100 GDPR; *Heilmann/Schulz*, in: Gierschmann/Schlender/Stentzel, Art. 40, cip. 1-5, and Art. 42, cip. 1-8.
[273] The DPA must approve the code of conduct (Art. 40 sect. 5 sent. 2 in combination with Art. 58 sect. 3 lit. d GDPR) or certification scheme (Art. 42 sect. 5 alt. 1 in combination with Art. 58 sect. 3 lit. f alt. 2 GDPR) and is, therefore, often already involved in the specification process, see, for example, Heilmann/Schulz, Art. 40, cip. 17, as well as Art. 41, cip. 5.
[274] See, regarding codes of conduct, Art. 41 sect. 4 as well as 5 GDPR, and regarding certification schemes, Art. 42 sect. 5 alt. 1, sect. 6 and 7 sent. 2 alt. 1 GDPR, as well as Art. 43 GDPR).
[275] See *Paal*, in: Paal/Pauly, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 40 cip. 3; *von Braunmühl*, in: Plath, BDSG/DSGVO, 3. Aufl., 2018, Art. 40 cip. 2.



3.2.2 Common advantages

These co-regulatory instruments promise several advantages. The main advantage is that such a co-regulatory strategy makes it possible to specify the broad and vague provisions of the GDPR with respect to the specific processing context. As protection measures can be more precisely adapted to the specific risks caused by the data processing in question, such a co-regulatory strategy can therefore lead, on the one hand, to a more effective protection of the data subjects. Codes of conduct and certification schemes can thus increase citizens' confidence in products and/or services based on data relating to them (in the following "data-driven products"). On the other hand, such an adaptation of the protection measures to the specific risks can at the same time avoid over-regulation, which in turn can lead to greater acceptance of compliance by those subject to regulation.[276]

In addition, the approval of the code of conduct and certification scheme by the competent DPA also increases legal certainty:

- Data controllers can use their adherence to a certification scheme as an element to demonstrate compliance with the privacy by design requirement[277]
- Data processors can use adherence to a code of conduct and the certification scheme to demonstrate the necessary guarantees[278]
- Data controllers and processors can also adhere to a code of conduct or certification scheme to

[276] See *Heilmann/Schulz*, Art. 43 cip. 1-8, in: Schlender/Stentzel/Veil Kommentar zur DS-GVO, 2nd ed., 2020, the latter comment also with respect to further advantages in favour of the executive and judiciary since often private certification bodies and associations carry most of the control resources Similarly, taking the aspect of legal certainty into consideration: *Bergt*, in: Kühling/Buchner, Datenschutz-Grundverordnung, 2. Aufl., 2018, Art. 40, cip. 1; *von Braunmühl*, in: Plath Art. 40 cip. 2
[277] See Art. 25 sect. 3 GDPR.
[278] See Art. 28 sect. 5 GDPR.

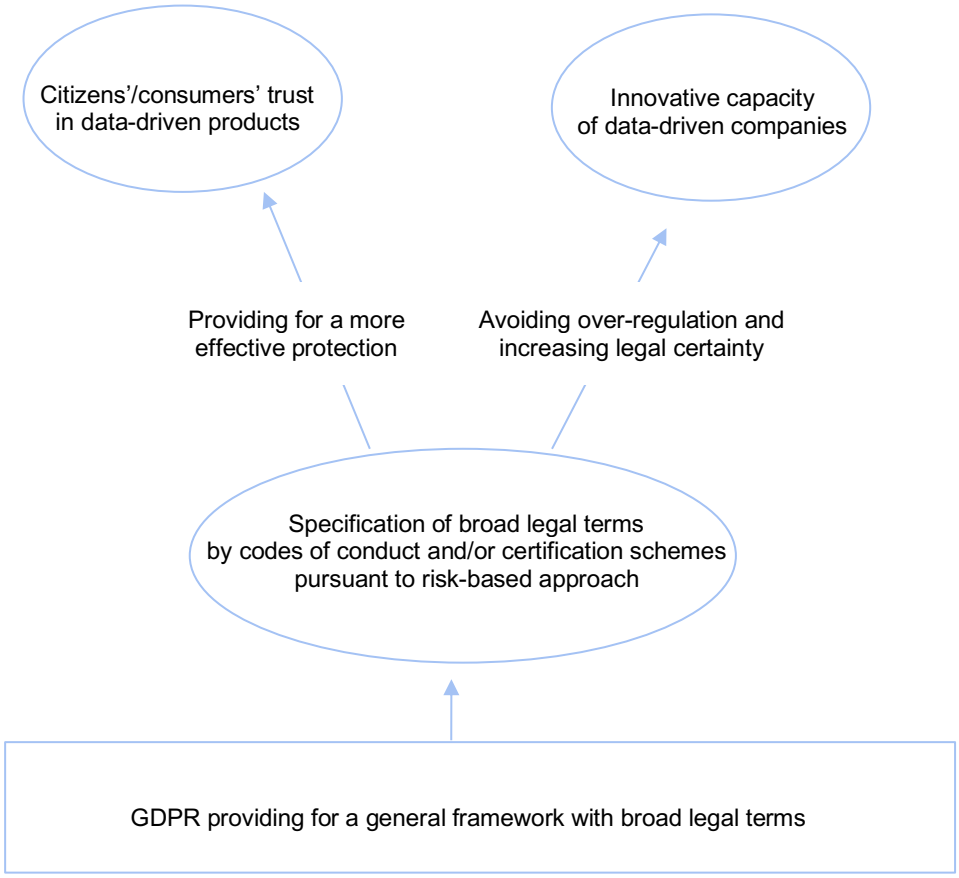
show compliance with the security by design requirement[279]

- Codes of conduct and certification schemes can also serve as a basis for the data transfer to a third country[280]

- Controllers and processors can refer to a code of conduct in the course of a data protection impact assessment (as highlighted in the DPIA, this can be particularly useful for micro, small and medium-sized companies that would like to use the collected data and are required to conduct an additional and even more specific DPIA)[281]

- Compliance with a code of conduct or a certification scheme can also play a role in decisions to impose a fine or the amount of such a fine.[282]

The following picture illustrates this idea:



In summary, it should be noted that the concept applied in the law regarding the co-regulatory instruments has an inherent "regulatory conflict". On the one hand, the more specific the code of conduct or certification scheme is, the more likely all these legal privileges are to apply. The reason for this is that these legal privileges are justified only if they actually meet the regulatory aim, i.e. to increase

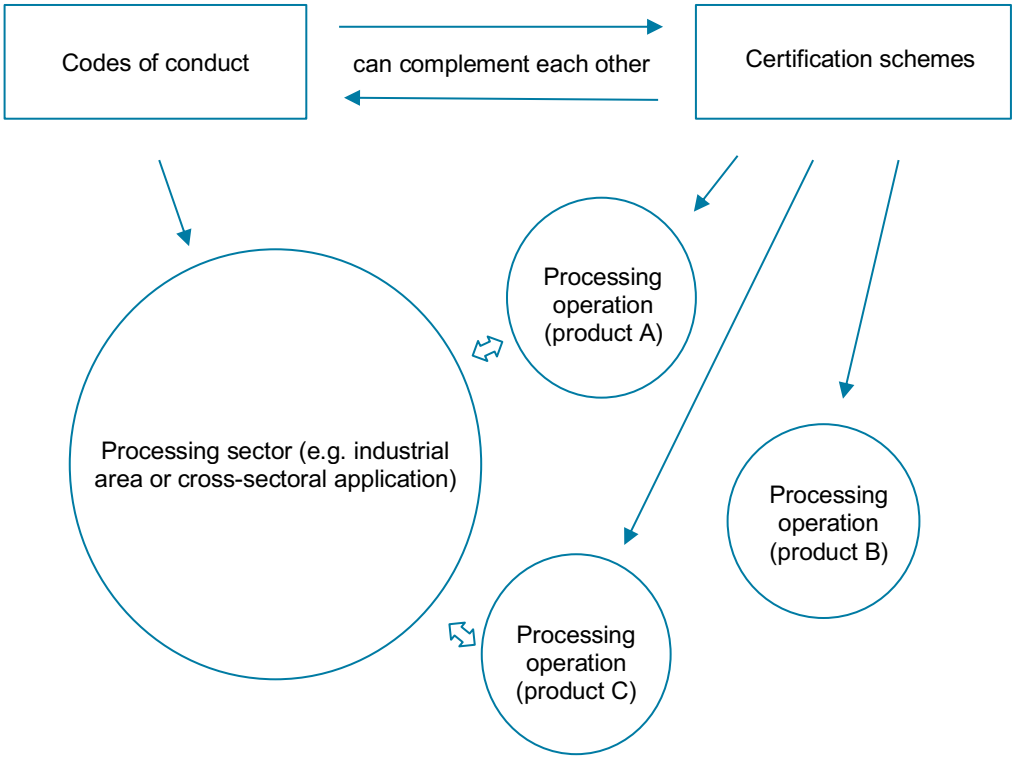
[279] See Art. 32 sect. 3 GDPR.
[280] See Art. 46 sect. 2 lit. e) and f) GDPR; this can also be guaranteed by binding corporate rules (BCR). However, because BCR only provide for this specific advantage, but not the others listed, BCR do not play a further role in this study.
[281] See Art. 35 sect. 8 GDPR; see DPIA, "I. 2.2.3 Legal and further implications".
[282] See Art. 83 sect. 2 sent. 2 lit. j) GDPR.

legal certainty by specifying broad and vague legal terms.[283] On the other hand, the more specific a code of conduct or certification schemes becomes, the more likely it is to limit the room for manoeuvre of the regulatory addressee and to increase the regulatory burden. This creates a regulatory conflict, as the GDPR expressly states that the needs of micro, small, and medium-sized companies must be taken into account when drawing up a code of conduct or when specifying the criteria of a certificate.[284] Thus, the risk-based approach can also play an essential role.[285]

3.2.3 Differences

Last but not least, keeping these advantages of both codes of conduct and certification schemes in mind, there are also certain differences. For example, the certification of a data-driven product, as opposed codes of conduct, may give a competitive advantage to users (of such a certificate), as it may signal to consumers a certain level of data protection that they may prefer to competitors' products..[286] More relevant for this study, however, is the fact that although codes of conduct refer to a certain processing sector, certification schemes refer to specific processing operations. Their object of evaluation is therefore different. This makes it possible to combine the two instruments in a complementary way to establish the most effective control mechanism.[287] The following picture illustrates this situation:

[283] See, for example, the wording of recital 98 sent. 1 GDPR: "Associations or other bodies representing categories of controllers or processors should be encouraged to draw up codes of conduct, within the limits of this Regulation, so as to facilitate the effective application of this Regulation, taking account of the specific characteristics of the processing carried out in certain sectors and the specific needs of micro, small and medium enterprises."
[284] See Art. 40 sect. 1 and 42 sect. 1 sent. 2 GDPR.
[285] See, for example, the wording of recital 89 sent. 2 GDPR: "In particular, such codes of conduct could calibrate the obligations of controllers and processors, taking into account the risk likely to result from the processing for the rights and freedoms of natural persons."; cf. *Heilmann/Schulz*, Art. 42, cip 35, in: Schlender/Stentzel/Veil Kommentar zur DS-GVO, 2nd ed., 2020
[286] See Grafenstein M., Co-Regulation and the Competitive Advantage in the GDPR: Data Protection Certification Mechanisms, Codes of Conduct and the "State of the Art" of Data Protection-by-Design' in G González-Fuster, R van Brakel and P De Hert (eds), *Research Handbook on Privacy and Data Protection Law: Values, Norms and Global Politics, Edward Elgar Publishing*. Edward Elgar Publishing
[287] See *Heilmann/Schulz*, Art. 42 cip. 10, in: Schlender/Stentzel/Veil Kommentar zur DS-GVO, 2nd ed., 2020; *von Braunmühl*, in: Plath, BDSG/DSGVO, 3. Aufl., 2018, Art. 42 cip. 5.



3.3 Conclusion: Whose interests are represented in the steering boards?

In conclusion, this DPIA could elaborate several important aspects not only for an appropriate data protection by design strategy for the intended system, but even methodically. First of all, readers will find that the risk assessment is somewhat repetitive since several arguments have been presented previously. A first assessment actually starts with the question in which sense the data relate to individuals so that data protection laws apply. Even if this DPIA concludes that all collected data in the intended system is personal, it makes sense to differentiate between the different types of collected and retrieved information and different forms of how this information relates to individuals to define the appropriate data protection by design strategy. A second assessment has been conducted with respect to the balancing test of the "legitimate interests"-clause according to Art. 6 sect. 1 lit. f GDPR. This assessment highlights the importance of the following protection measures: first, to pseudonymise the data to the greatest extent, for example, by renewing the hashes of the movement patterns of data subjects at sufficiently short intervals (or according to specific location areas); second, to control how the data is used later on; and third, to make the intended processing as transparent as possible. A further assessment was made when examining more closely the requirements of the principle of purpose limitation with particular respect to research and statistical purposes (such as in an urban traffic management environment). In this respect, it became clear that the intended system should not only apply a decentralised infrastructure. More importantly, this system does not at all imply that the data is stored for all kinds of research purposes in advance. Rather, the system may simply consist of a predefined set of technical, organisational and legal standards that will make it much easier to collect, combine and process data for research purposes when these purposes actually require it. Thus, several main building blocks for the technical and organisational design of the intended system have already been defined before the actual risk assessment could take place.

However, following a methodological refinement of the risk assessment, it was possible to identify the main impact of the intended system on specific data subjects' fundamental rights and thus to create a structure to prioritise and adjust the technical-organisational protection measures accordingly. On this basis, it became definitely clear that the ultimate challenge of the system is how to ensure the correct application of the principles at a later stage, i.e. when the data are used. To meet this challenge, codes of conduct and certification schemes can play a major role, as they provide institutionalised mechanisms that guarantee control over time. As a result, this assessment could not go into detail on the sources of risk, threats and measures, as the objective of the evaluation could not yet be sufficiently defined. But the risk assessment could help in deciding what measures to focus on when defining the appropriate data protection by design strategy. Of course, many (more detailed) questions remain unanswered. One question certainly is which stakeholders act as a controller, processor or as joint controllers and therefore which of them has what responsibility Another question is to what extent it is possible to determine the specific risks and, accordingly, the protective measures before the system is actually built – and which (yet unspecified) risks and appropriate protection measures must be specified later one. A third important question is who will sit on the steering committee when certification and/or monitoring bodies will carry out these specification tasks. Will these be only the (representatives of) the controllers and processors in the envisaged system or also (from) the data subjects?

REFERENCES

AK Technik der Konferenz der unabhängigen Datenschutzbehörden des Bundes und der Länder (ed.). (2018). *Das Standard- Datenschutzmodell: Eine Methode zur Datenschutzberatung und -prüfung auf der Basis einheitlicher Gewährleistungsziele*. Standard-Datenschutzmodell. https://www.datenschutzzentrum.de/uploads/sdm/SDM-Methode_V1.1.pdf

Albrecht, J. P., & Jotzo, F. (2017). *Das neue Datenschutzrecht in der EU: Grundlagen, Gesetzgebungsverfahren, Synopse*. Nomos.

Bieker F, (2018). Die Risikoanalyse Nach Dem Neuen EU-Datenschutzrecht Und Dem Standard-Datenschutzmodell, Datenschutz und Datensicherheit, 42(1), 27-31

Bieker, F., Bremert, B., & Hansen, M. (2018). Die Risikobeurteilung nach der DSGVO. *Datenschutz und Datensicherheit*, *42*(8), 492–496. https://doi.org/10.1007/s11623-018-0986-1

Black, J. (2008). Forms and Paradoxes of Principles Based Regulation. *London School of Economics and Political Science*, *13*. www.lse.ac.uk/collections/law/wps/wps.htm

Bohnert, J., Krenberger, B., & Krumm, C. (ed.). (2016). *Ordnungswidrigkeitengesetz* (4. Ed.). C.H.Beck.

Brink, S., & Eckhardt, J. (2015). Wann ist ein Datum ein personenbezogenes Datum? Anwendungsbereich des Datenschutzrechts. *ZD*, *5*, 205.

Britz G, (2010). Informationelle Selbstbestimmung Zwischen Rechtswissenschaftlicher Grundsatzkritik Und Beharren Des Bundesverfassungsgerichts, Offene Rechtswissenschaft 561

Buchner, B. (2006). *Informationelle Selbstbestimmung im Privatrecht*. Mohr Siebeck.

Censky, A. (2011, November 22). *Malls track shoppers' cell phones on Black Friday*. CNNMoney. https://money.cnn.com/2011/11/22/technology/malls_track_cell_phones_black_friday/index.htm

Cleverciti Sensor. (2020, January 1). *Realtime, accurate detection of available parking spaces*. https://www.cleverciti.com/wp-content/uploads/2020/01/cleverciti_Sensor_1page_WEB.pdf

Commission Nationale Informatique & Libertés. (2018). *Privacy Impact Assessment (PIA)*. https://www.cnil.fr/sites/default/files/atoms/files/cnil-pia-2-en-templates.pdf

Contreras, J. L., & Reichman, J. H. (2015). Sharing by design: Data and decentralized commons. *Science*, *350*(6266), 1312. https://doi.org/10.1126/science.aaa7485.

Corrales, M., Fenwick, M., & Forgo, N. (2017). *New Technology, Big Data and the Law*. Springer.

Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679. Accessed 17. March 2020, von https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611236

Czernik, A. (2015, September 4). *IP-Adressen – Funktion, Aufbau, Tracking*. Datenschutzbeauftragter Info. https://www.datenschutzbeauftragter-info.de/ip-adressen-funktion-aufbau-tracking/

Drucker, P. (2002). The Discipline of Innovation. *Harvard Business Review*, 95.

Duncan, G. T., Elliot, M., & Salazar-González, J.-J. (2011). *Statistical Confidentiality: Principles and Practice*. Springer.

Ehmann, E., & Selmayr, M. (2018). *Datenschutz-Grundverordnung: DS-GVO* (2. Ed.). C.H.Beck.

Eisele, D., Grigorjew, O., Karaboga, M., Matzner, T., Morlock, T., Nebel, M., Ochs, C., Rohbrahn, R., Rzepka, C., & Fhom, H. (2017). *Privatheit in öffentlichen WLANs Spannungsverhältnisse zwischen gesellschaftlicher Verantwortung, ökonomischen Interessen und rechtlichen Anforderungen* (White Paper). Forum Privatheit.

https://www.forum-privatheit.de/privatheit-in-oeffentlichen-wlans/

El Emam, K., & Alvarez, C. (2015). A critical appraisal of the Article 29 Working Party Opinion 05/2014 on data anonymization techniques. *International Data Privacy Law*, *5*(1), 73–87. https://doi.org/10.1093/idpl/ipu033

Elliot, M., Mackey, E., O'Hara, K., & Tudor, C. (2016). The Anonymisation Decision-Making Framework. *UK Anonymisation Network*. https://ukanon.net/wp-content/uploads/2015/05/The-Anonymisation-Decision-making-Framework.pdf

Engeler, M., & Felber, W. (2017). Entwurf der ePrivacy-VO aus Perspektive der aufsichtsbehördlichen Praxis: Reguliert der Entwurf an der technischen Realität vorbei? *ZD*, *6*, 251.

European Data Protection Board (EDPB) having adopted Article 29 Data Protection Working Party. (2007). *Opinion 4/2007 on the concept of personal data* (01248/07/EN WP 136). European Union. https://www.clinicalstudydatarequest.com/Documents/Privacy-European-guidance.pdf

European Data Protection Board (EDPB) having adopted Article 29 Data Protection Working Party. (2013). *Opinion 03/2013 on purpose limitation* (00569/13/EN WP 203). European Union. https://www.bfdi.bund.de/SharedDocs/Publikationen/DokumenteArt29Gruppe_EDSA/Stellungnahmen/WP203_Opinion22013PurposeLimitation.html

European Data Protection Board (EDPB) having adopted Article 29 Data Protection Working Party. (2014). *Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC* (844/14/EN WP 217; S. Brussels). European Union. http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp217_en.pdf

European Data Protection Board (EDPB) having adopted Art. 29 Data Protection Working Party, (2017) '*Guidelines on Data Protection Impact Assessment (DPIA) and Determining Whether Processing Is "Likely to Result in a High Risk" for the Purposes of Regulation 2016/679*'. European Union. https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611236

European Data Protection Board (EDPB). (2018). *Statement of the EDPB on the revision of the ePrivacy Regulation and its impact on the protection of individuals with regard to the privacy and confidentiality of their communications*. European Union. https://edpb.europa.eu/our-work-tools/our-documents/drugi/statement-edpb-revision-eprivacy-regulation-and-its-impact_en

European Union Agency for Fundamental Rights, & Council of Europe. (2014). *Handbook on European data protection law*. Publications Office of the European Union. http://fra.europa.eu/en/publication/2014/handbook-european-data-protection-law

Exploring the 'Design' in Privacy by Design. (2018, February 6). Computer Privacy and Data Protection Conference 2018. Panel discussion accessible via Youtube. https://www.youtube.com/watch?v=qMwQRB7nJwo

Fisher, D. (2014, March 19). *Research Finds MAC Address Hashing Not a Fix for Privacy Problems*. ThreatPost. https://threatpost.com/research-finds-mac-address-hashing-not-a-fix-for-privacy-problems/104893/

Friedewald, M., Obersteller, H., Nebel, M., Bieker, F., & Rost, M. (2017). *Datenschutz-Folgenabschätzung. Ein Werkzeug für einen besseren Datenschutz* (Nr. 3; White Paper). Forum Privatheit. https://www.forum-privatheit.de/publikationen/white-paper-policy-paper/

Fueglistaller, U., Müller, C. A., Müller, S., & Volery, T. (2016). *Entrepreneurship: Modelle – Umsetzung – Perspektiven. Mit Fallbeispielen aus Deutschland, Österreich und der Schweiz* (4. Ed.). Springer Gabler.

Gagzow, G. (2019). *Deliverable 3.2: Datenschutz bei vernetzten, automatisierten und kooperativen Fahrzeugen nach der Datenschutzgrundverordnung*. Integrierte Kommunikationsplattform für automatisierte Elektrofahrzeuge. https://www.datenschutzzentrum.de/uploads/projekte/ikopa/iKoPA_D3.2-3.pdf

Gierschmann, S., & Saeugling, M. (ed.). (2014). *Systematischer Praxiskommentar Datenschutzrecht. Datenschutz aus Unternehmenssicht*. Bundesanzeiger Verlag.

Gierschmann, S., Schlender, K., Stentzel, R., & Veil, W. (2020). *Kommentar zur Datenschutz-Grundverordnung* (2. Ed.). Reguvis.

Gola, P. (ed.). (2018). *Datenschutz-Grundverordnung: DS-GVO VO (EU) 2016/679* (2. Ed.). C.H. Beck.

Gola, P., Heckmann, D., Braun, F., & Germany (ed.). (2019). *Bundesdatenschutzgesetz: BDSG* (13. Ed.). C.H. Beck.

Graf, J. (2019). *BeckOK OWiG* (24. Ed.). C.H. Beck.

Grafenstein, M. von. (2018). *The Principle of Purpose Limitation in Data Protection Laws: The Risk-Based Approach, Principles, and Private Standards as Elements for Regulating Innovation* (Bd. 12). Nomos.

Grafenstein M., (2019). Co-Regulation and the Competitive Advantage in the GDPR: Data Protection Certification Mechanisms, Codes of Conduct and the "State of the Art" of Data Protection-by-Design' in G González-Fuster, R van Brakel and P De Hert (eds), *Research Handbook on Privacy and Data Protection Law: Values, Norms and Global Politics, Edward Elgar Publishing*. Edward Elgar Publishing

Grafenstein, M. V., & Schulz, W. (2015). The right to be forgotten in data protection law: A search for the concept of protection. *International Journal of Public Law and Policy*, *5*(3), 249. https://doi.org/10.1504/IJPLAP.2015.075049

Grafenstein, M. von, & Wernick, A. (o. J.). Data Governance. *HIIG*. Accessed 17. March 2020, von https://www.hiig.de/en/project/data-governance/

Härtler, G. (2014). *Statistisch gesichert und trotzdem falsch? Vom (Un-)Wesen statistischer Schlüsse*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-43357-7

Hoffmann-Riem, W. (2010). *Offene Rechtswissenschaft: Ausgewählte Schriften von Wolfgang Hoffmann-Riem und begleitende Analysen*. Mohr Siebeck.

Hoffmann-Riem, W., Schmidt-Aßmann, E., & Voßkuhle, A. (ed.). (2012). *Grundlagen des Verwaltungsrechts: Methoden, Maßstäbe, Aufgaben, Organisation* (2. Aufl., Bd. 1). Beck.

Kühling, J., Buchner, B., Bäcker, M., Bergt, M., Boehm, F., Caspar, J., & Dix, A. (2018). *Datenschutz-Grundverordnung/BDSG: Kommentar* (2. Auflage). C.H. Beck.

Kühling, J., & Martini, M. (2016). *Die Datenschutz-Grundverordnung und das nationale Recht: Erste Überlegungen zum innerstaatlichen Regelungsbedarf*. Monsenstein und Vannerdat.

Kühn, U. (2009). Geolokalisierung mit anonymisierten IP-Adressen. *Datenschutz und Datensicherheit - DuD*, *33*(12), 747–751. https://doi.org/10.1007/s11623-009-0196-y

Kurzpapier Nr. 18: Risiko für die Rechte und Freiheiten natürlicher Personen. (2018, April 26). Datenschutzkonferenz. https://www.datenschutzkonferenz-online.de/media/kp/dsk_kpnr_18.pdf

Mitsch, W. (ed.). (2018). *Karlsruher Kommentar zum Gesetz über Ordnungswidrigkeiten: OWiG* (5. Ed.). C.H. Beck.

Moroz, P., & Hindle, K. (2012). Entrepreneurship as a Process: Toward Harmonizing Multiple Perspectives. *Entrepreneurship Theory and Practice*, *7*, 781. https://doi.org/10.1111/j.1540-6520.2011.00452.x

Ohm, P. (2009). Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. *UCLA Law Review*, *57*, 1701.

Opt out of Smart Store tracking. (o. J.). Future of Privacy Forum. Accessed 17. March 2020, von https://optout.smart-places.org

Paal, B., & Pauly, D. (2018). *Datenschutz-Grundverordnung Bundesdatenschutzgesetz: DS-GVO BDSG* (2. Ed.). C.H. Beck.

Peers, S., Hervey, T., Kenner, J., & Ward, A. (ed.). (2014). *The EU Charter of Fundamental Rights*. C.H. Beck, Hart Publishing, Nomos.

Plath, K.-U., Becker, T., Krohm, N., Braunmühl, P. von, Kuhnke, M., Bussche, A. von dem, Grages, J.-M., Roggenkamp, J. D., Hullen, N., Schreiber, L., Jenny, V., Stamer, P., Kamlah, W., & Wittmann, J. (2018). *DSGVO/BDSG: Kommentar zu DSGVO, BDSG und den Datenschutzbestimmungen von TMG und TKG* (3. Auflage). Otto Schmidt.

Posner, R. A. (2011). *Economic Analysis of Law* (8. Ed.). Wolters Kluwer Law & Business: Aspen Publishers.

Privacy by design in smart cities. (o. J.). *HIIG*. Accessed 6. Mai 2020, von https://www.hiig.de/en/project/privacy-by-design-in-smart-cities/

Privacy protector. (o. J.). Accessed 14. Mai 2020, von https://www.european-privacy-seal.eu/EPS-en/privacy-protector

Proposal for a Regulation of the European Parliament and of the Council concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications), Pub. L. No. COM/2017/010 final-2017/03 (COD), 13.20.60.00. Accessed 17. March 2020, von https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52017PC0010

Recital 92—Broader data protection impact assessment. (o. J.). *General Data Protection Regulation (GDPR)*. Accessed 17. March 2020, von https://gdpr-info.eu/recitals/no-92/

Registration Authority. (o. J.). IEEE Standards Association. Accessed 17. March 2020, von https://standards.ieee.org/products-services/regauth/oui/index.html

Roßnagel, A. (1993). *Rechtswissenschaftliche Technikfolgenforschung: Umrisse einer Forschungsdisziplin* (1. Auflage). Nomos.

Rost, M. (2012). Standardisierte Datenschutzmodellierung. *Datenschutz und Datensicherheit*, *36*(6), 433–438. https://doi.org/10.1007/s11623-012-0153-z

Schumpeter, J. A. (2008). *Capitalism, socialism and democracy* (3. Ed.). Harper Perennial Modern Thought.

Simitis, S., Hornung, G., & Spiecker gen. Döhmann, I. (ed.). (2019). *Datenschutzrecht. DSGVO mit BDSG*. Nomos.

Staben, J. (2016). *Der Abschreckungseffekt auf die Grundrechtsausübung: Strukturen eines verfassungsrechtlichen Arguments*. Mohr Siebeck.

Stoklas, J., & Wendorf, J. (2017). Der Staatstrojaner – verhältnismäßig unverhältnismäßig? *ZD-Aktuell*, *14*.

TU Dresden. (o. J.). *Projekt Mobilität in Städten SRV*. Accessed 17. March 2020, von https://tu-dresden.de/bu/verkehr/ivs/import/vip/srv

Unabhängige Landeszentrum für Datenschutz Schleswig-Holstein, & European Union Agency for Cybersecurity. (2019, November 12). *ULD - ENISA Workshop: Pseudonymisation and relevant security technologies* [Event]. European Union Agency for Cybersecurity. https://www.enisa.europa.eu/events/uld-enisa-workshop/uld-enisa-workshop-pseudonymization-and-relevant-security-technologies

Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein. (2009). *EuroPriSe-Siegel für datenschutzfreundliche Videoüberwachungssoftware „Privacy Protector"* [Pressemitteilung]. European Privacy Seal. https://www.european-privacy-seal.eu/EPS-en/privacy-protector

van Brakel, R., & De Hert, P. (ed.). (2018). *Understanding American Privacy: Research Handbook on Privacy and Data Protection Law: Values, Norms and Global Politics*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3256918

van Dijk, N., Gellert, R., & Rommetveit, K. (2016). A risk to a right? Beyond data protection risk assessments. *Computer Law & Security Review*, *32*(2), 286–306. https://doi.org/10.1016/j.clsr.2015.12.017

Vanhoef, M., Matte, C., Cunche, M., Cardoso, L. S., & Piessens, F. (2016). Why mac address randomization is not enough: An analysis of wi-fi network discovery mechanisms. *Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security - ASIA CCS '16*, 413–424. https://doi.org/10.1145/2897845.2897883

Verkehrslenkung Berlin - VLB C. (2014). *Straßenverkehrszählung Berlin 2014* [Ergebnisbericht]. Senatsverwaltung für Stadtentwicklung und Umwelt. https://www.berlin.de/senuvk/verkehr/lenkung/vlb/download/Ergebnisbericht_2014.pdf

Wolff, A., & Brink, S. (2017). *Beck'scher Online-Kommentar Datenschutzrecht* (19. Ed.). C.H. Beck.