HAL open science

# Data Protection Impact Assessment: A Hands-On Tour of the GDPR's Most Practical Tool

Felix Bieker, Nicholas Martin, Michael Friedewald, Marit Hansen

## To cite this version:

Felix Bieker, Nicholas Martin, Michael Friedewald, Marit Hansen. Data Protection Impact Assessment: A Hands-On Tour of the GDPR's Most Practical Tool. Marit Hansen; Eleni Kosta; Igor Nai-Fovino; Simone Fischer-Hübner. Privacy and Identity Management. The Smart Revolution : 12th IFIP WG 9.2, 9.5, 9.6/11.7, 11.6/SIG 9.2.2 International Summer School, Ispra, Italy, September 4-8, 2017, Revised Selected Papers, AICT-526, Springer International Publishing, pp.207-220, 2018, IFIP Advances in Information and Communication Technology, 978-3-319-92924-8. 10.1007/978-3-319-92925-5_13 . hal-01883612

**HAL Id: hal-01883612**

**https://inria.hal.science/hal-01883612v1**

Submitted on 28 Sep 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

# Data Protection Impact Assessment: A Hands-On Tour of the GDPR's Most Practical Tool

Felix Bieker[1], Nicholas Martin[2], Michael Friedewald[2], Marit Hansen[1]

[1] Unabhängiges Landeszentrum für Datenschutz (ULD, Independent Centre for Privacy Protection) Schleswig-Holstein, Kiel, Germany
{fbieker|marit.hansen}@datenschutzzentrum.de

[2] Fraunhofer Institute for Systems and Innovation Research ISI, Karlsruhe, Germany
{nicholas.martin|michael.friedewald}@isi.fraunhofer.de

**Abstract.** This workshop introduced participants to the process of Data Protection Impact Assessment. This new tool of the GDPR is highly relevant for any processing of personal data, as it helps to structure the process, be aware of data protection issues and the relevant legislation and implement proper safeguards to protect data subjects. For processing operations posing a high risk for data subjects, a DPIA is mandatory from May 2018. The interactive workshop provided a framework for DPIA and guidance on specific questions such as when a high risk is likely to occur or how specific risks can be evaluated, which was assessed by participants in an interactive session with two different scenarios.

**Keywords:** Data Protection Impact Assessment, Risk to Rights and Freedoms, General Data Protection Regulation, Data Protection, EU law

## 1 Introduction

The General Data Protection Regulation (GDPR) will replace the Data Protection Directive on 25 May 2018. Among the regulatory and governance instruments it introduces is the Data Protection Impact Assessment (DPIA), which serves to mitigate risks to the rights and freedoms of natural persons and is a tool for controllers to conform to the GDPR's legal requirements. DPIA builds on Privacy Impact Assessments (PIAs), as they have been encouraged by academia [1], Data Protection Authorities (DPAs) [2] and the European Commission (e. g. for RFID applications [3]). However, DPIA focuses on conformity to EU data protection law and thus has a more specific scope. It is a very useful tool for controllers to control their processing of personal data and ensure compliance.

When a high risk to the rights of individuals is likely, carrying out a DPIA is mandatory according to Article 35(1) GDPR. While non-compliance with this obligation may incur a penalty of up to 2% of the world-wide annual turnover of a business according to Article 83(4)(a) GDPR, the notion of high risk is not defined in the Regulation. Rather, Article 35(3) GDPR lists a few examples of data processing operations,

which could potentially pose a high risk. Similarly, the GDPR does not offer much advice about how to carry out a DPIA; much less a methodology. Article 35(4) GDPR contains only minimal requirements, and provides no further guidance about how to implement these in practice. Furthermore, existing processes for Privacy Impact Assessments (PIA) may not take due account of the GDPR's legal requirements, such as data protection by design and by default, which is now enshrined in Article 25 GDPR, or the risk-based approach adopted in this new legislation.

Thus, the goal of the workshop was to acquaint participants with the DPIA framework, how it can best be carried out and which specific issues may arise. Participants were first introduced to the DPIA framework developed by the German research consortium *Privacy Forum* (Forum Privatheit) [4, 5] and focuses on the rights of individuals. This framework is based on the legal requirements of the upcoming GDPR, in particular Article 35, as well as the Standard Data Protection Model (SDM) methodology adopted by the German data protection authorities [6], which operationalises these legal requirements, and best practices. The framework takes account of the *Guidelines on Data Protection Impact Assessment* of the Article 29 Data Protection Working Party [7]. In order to raise the participants' awareness of the risks to the rights and freedoms of individuals two case studies were discussed with a view to identifying the relevant risks by applying the data protection goals systematised in the SDM.

## 2 Introduction to Data Protection Impact Assessments

A DPIA begins before any data are processed and continues throughout the life cycle of a project and its data processing operations. It is a useful tool for any controller to implement their obligations under the GDPR and allows them to document this, as they are obliged to under Article 5(2) GDPR. At the heart of this process is the analysis of risks to the rights and freedoms of individuals that may emanate from the processing of personal data and is the basis for mitigating these risks through technical and organisational measures. This can best be achieved in four phases, as detailed in Figure 1 below.

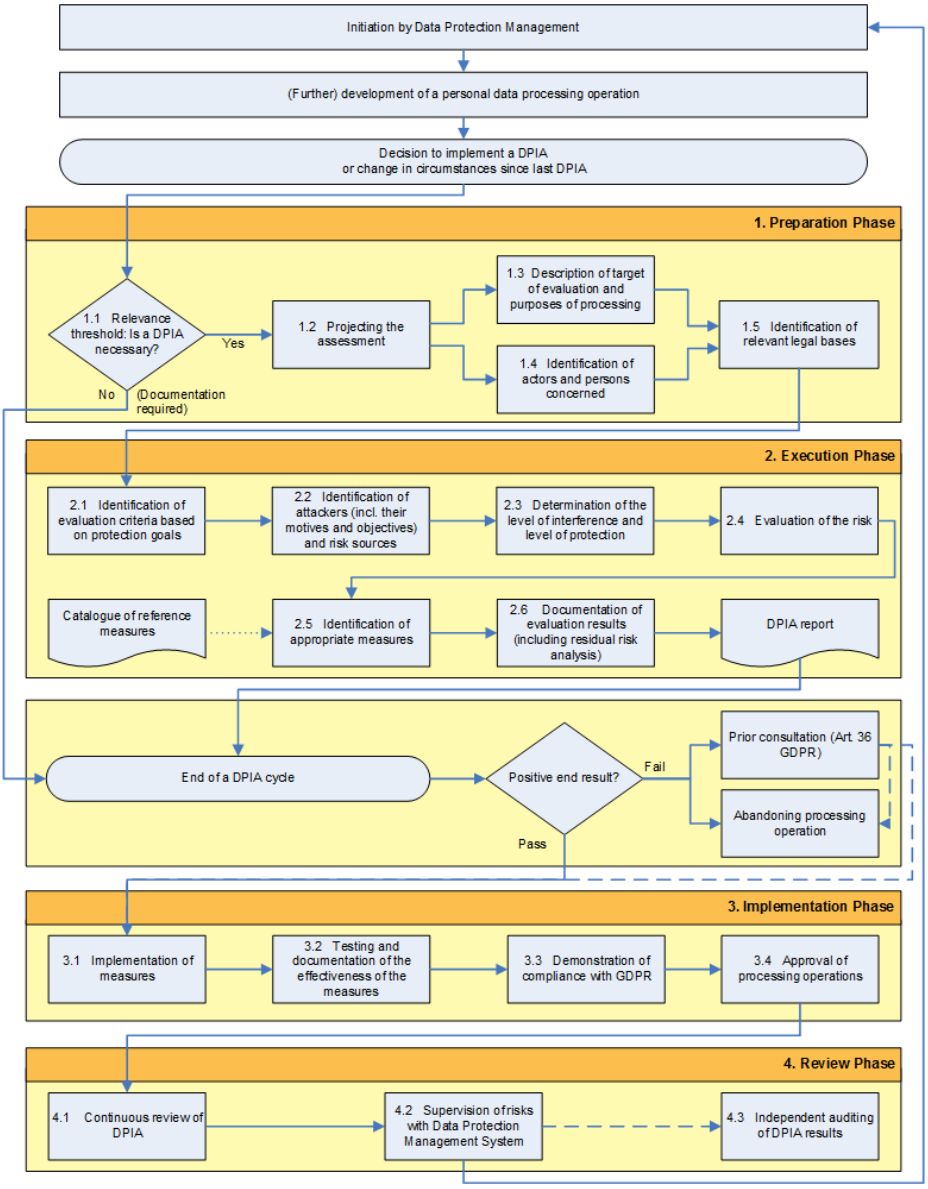


**Fig. 1.** Framework for Data Protection Impact Assessment

In the preparation phase (1.), a team is assembled to carry out the assessment and relevant information about the envisaged processing collected. In the execution phase (2.) the sources of risk (i.e. attackers) are identified, the gravity of the interference is determined and the risks for the rights and freedoms of natural persons are evaluated. Furthermore, the controller identifies appropriate measures and documents the results

of the evaluation in a DPIA report. On the basis of this evaluation, the controller then decides whether to carry out the envisaged processing operation or not. If the DPIA finds that the risks to the rights of individuals remain high even with the identified measures, the controller has to consult the supervisory authority according to Article 36 GDPR before the processing can start. The controller may also decide to abandon the processing operation.

If adequate measures could be identified to address the risks and ensure the protection of the rights of individuals (or this is achieved during the consultation with the supervisory authority), the controller implements these measures, tests and documents their effectiveness and demonstrates compliance with the GDPR (3.), before approving the processing operation. In the review phase (4.) the controller monitors the risks for the rights and freedoms of natural persons and repeats (parts of) the assessment when necessary.

### 2.1 The Standard Data Protection Model and Risk Analysis

The Standard Data Protection Model is a methodology to ensure effective compliance with data protection obligations and allows for auditing and control through transparent processes. This is achieved by formulating explicit data protection goals, which are derived from the legal requirements of data protection law. The data protection goals are the following:

**Data Minimisation.** Data minimisation substantiates and operationalises the principle of necessity, which requires of any process as a whole as well as any of its steps not to collect, process and use more personal data than necessary for the achievement of the purpose of the processing. Data minimisation is to be taken into account proactively as an element of data protection-friendly design. Starting with the design of information technology by the manufacturer and its configuration and adaptation to the operating conditions, to its use in the core and auxiliary processes of the operation, for instance in the maintenance of the systems used; from the collection of personal data, through its processing and use, to its erasure or complete anonymization; throughout the entire life cycle of the data.

**(1) Availability.** Personal data must be available and can be used properly in the intended process. Thus, the data must be accessible to authorised parties and the methods intended for their processing must be applied. This presupposes that the methods can deal with the available data formats. Availability comprises the ability to find specific data (e.g. by means of address directories, reference or file numbers), the ability of the employed technical systems to make data accessible to individuals in an adequate manner, and the possibility to interpret the content of the data (semantic ascertainability).

**(2) Integrity.** On the one hand, information technology processes and systems must continuously comply with the specifications that have been determined for the execution of their intended functions. On the other hand, integrity means that the data to be processed remain intact, complete, and up-to-date. Deviations from these properties must be excluded or at least be ascertainable so that this can either be taken into consideration or the data can be corrected. If the protection goal integrity is understood as a form of accuracy within the meaning of Article 5(1)(d) GDPR, this leads to the claim that there is sufficient congruency between the legal-normative requirement and common practice, both in terms of technical detail as well as in the broad context of the procedure and its overall purpose.

**(3) Confidentiality.** No person is allowed to access personal data without authorisation. A person is not only unauthorised when it is a third party external to the controller, regardless of whether they act with or without a criminal intent, but also employees of technical service providers who do not need access to personal data for the provision of the service, or persons in organisational units who are unrelated to the respective procedure or data subject.

**(4) Unlinkability.** Data shall be processed and analysed only for the purpose for which they were collected. Data sets can in principle be processed for further purposes and can be combined with other, potentially publicly available data. Larger and more meaningful data sets also increase the potential for abuse, i.e. to use the data unlawfully, for purposes beyond the legal basis. Such further processing is lawful only in strictly defined circumstances. The GDPR only allows them to be used for archival purposes which are in the public interest, for scientific or historical research purposes or for statistical purposes, and explicitly calls for safeguards for the rights and freedoms of the data subjects. These safeguards are to be achieved through technical and organisational measures. In addition to measures of data minimisation and pseudonymisation, other measures that allow the further processing to be separated from the source processing are also suitable, ensuring separation both on the organisational and on the system side. The data base can, for example, be adapted to the new purpose by pseudonymisation or reduction of data volume.

**(5) Transparency.** The data subject as well as the system operators and the competent supervisory authorities must be able to understand, to a varying extent, which data are collected and processed for a particular purpose, which systems and processes are used for this purpose, where the data flow for which purpose, and who is legally responsible for the data and systems in the various phases of data processing. Transparency is necessary for the monitoring and control of data, processes, and systems from their origin to their erasure and is a prerequisite for lawful data processing. Informed consent, where it is necessary, can be given by data subjects only if these criteria are met. Transparency of the entire data processing operation and of the parties involved can help ensure that data subjects and supervisory authorities can identify deficiencies and, if necessary, demand appropriate procedural changes.

**(6) Intervenability.** Data subjects are effectively granted their rights to notification, information, rectification, blocking and erasure at any time, and that the controller is obliged to implement the appropriate measures. For this purpose, controllers must be able to intervene in the processing of data throughout the process; from the collection to the erasure of the data.

All of these protection goals can be linked to specific provisions in the GDPR, and all except availability, which is an implicit requirement throughout the GDPR, are mentioned in the principles relating to personal data processing in Article 5(1) [6].

Consequently, they can be used in the assessment of the risks to the rights and freedoms of natural persons in order to identify potential sources of risks and potential damages to these rights and freedoms. According to recital 76 the likelihood and severity of potential damages should be determined objectively and with reference to the nature, scope, context and purposes of the processing. However, as the phrasing of the risk already makes clear, recital 75 emphasises that this potential damage includes also non-material damage, such as discrimination, reputational damage, social disadvantages, the deprivation of data subjects' rights or preventing them from exercising control over their personal data. When read jointly with the second sentence of recital 94 it further becomes clear that besides such potential damages, interferences with fundamental rights, for instance the right to the protection of personal data under Article 8 CFR, the right to private life under Article 7 CFR, freedom of speech under Article 11 CFR or the right to be protected against discrimination under Article 21 CFR, are also risks to be considered in this assessment [8].

When considering this fundamental rights dimension of risk in the GDPR it also becomes clear that mathematical formulas, such as the common $R = \sum_{k=1}^{n} I_k \times p(I_k)$, where R is the risk, which is the product of the impact multiplied by the probability of potential damage, are not applicable here. Instead, the evaluation should classify the effect of potential damages or interferences with fundamental rights as well as the likelihood of their realisation into certain categories, such as marginal, limited, serious and severe. By applying each data protection goal to a processing operation to identify potential risks and then evaluating these, controllers can ensure that they fulfil their obligations with regard to the rights and freedoms of natural persons.

### 2.2 Data Subject Participation

As data protection law is ultimately concerned with safeguarding the rights of individuals, scientific studies have been demanding for years that DPIAs (or PIAs) should not only include the views of (technical and legal) experts. Making use of the expertise of technical experts alone may lead to a very narrow perspective (e.g. limiting an assessment to legal aspects alone) and also to a technocratic-paternalistic approach that takes decisions without taking citizens' concerns duly into account. Rather, a comprehensive and broad consultation of stakeholders is necessary to increase the quality of the assessment results and their legitimacy. However, the questions of which actors are considered to be 'relevant' at all and who determines this [9] should be kept in mind. This idea is reflected in Article 35(9) GDPR that stipulates that the

views of data subjects or their representatives on the intended processing should be taken into account – if appropriate. The provision limits the controller's obligation to allow participation with reference to the effort needed or other conflicting interests (security, intellectual property rights, etc.).

Without prejudice to such limitations, the question arises as to how the different stakeholder groups and interests can be involved in the evaluation process of a DPIA. In terms of methodology, relatively simple and proven methods are available, with which companies already have experience in the areas of product design and marketing (e.g. focus groups), but of course more elaborate methods from participatory Technology Assessment (pTA) can also be used [10]. However, an evaluation with data subject participation poses particular challenges in terms of timing and circumstances:

- The consultation should best start early in the process, so as to allow for an impact on the design of the processing operation.[1]
- The involvement of data subjects can be problematic, as careful and systematic assessment often requires expertise that lay people usually do not have. The key question here is therefore how this expertise can be conveyed to enable discussions on equal footing between lay people and experts.
- The vocabulary used in the evaluation process has implications for the intensity and quality of the involvement of different groups of actors. For example, certain forms of wording are likely to favour particularly technophile actors or those with legal knowledge. It will therefore be crucial to organise a neutral translation process between the different groups.

Extensive participatory DPIAs involving external stakeholders are, however, likely to remain the exception, since this process is time-consuming and could lead to consultation fatigue among certain stakeholder groups. Under normal circumstances the data subjects' views should be taken into account by involving units from the organisation that are in close and regular contact with the data subjects, i.e. sales, service or the works or staff council [11].

## 3 Hands-On: Practical Assessment of the Risks for the Rights and Freedoms of Natural Persons

After the input statements, participants were divided into two groups to discuss the two following case studies and identify risks for the rights and freedoms of natural persons. These were then summarized by participants of each group and discussed with all participants.

[1] However, when a DPIA is conducted prior to market launch or in parallel with the development process the inclusion of external persons may be undesirable. Reasons might be the immaturity of the technology, the organisation's desire for confidentiality or fear of a bad public image.

### 3.1 Case Studies

**Case Study 1: Smart Surveillance in Train Stations.**
After successful pilots, the national police force of an EU Member State has proposed setting up cameras with automated biometric recognition and behavioural analysis capabilities in all of the country's train stations. The system will have access to the images and biometric data from the national identity-card database, as well as police databases of terrorist and criminal suspects, political extremists and religious fanatics, and persons of interest or concern. It is supposed to be able to identify individuals with a very high degree of accuracy. The data will be stored for up to 1 year.

Besides identifying individuals, the system performs behavioural analysis to identify a range of suspicious behaviours (e.g., looking about a lot, avoiding security personnel, leaving luggage behind). It also identifies dangerous behaviour or behaviour indicating suicidal tendencies, especially of vulnerable individuals (e.g. drunken gait, straying close to platforms).

When the system picks up suspicious (or dangerous) behaviour or individuals, it sends automated messages to the station security personnel, city anti-terror units, and/or the station health and safety personnel (for drunks, etc.), whereupon these initiate enhanced monitoring or other interventions (e.g. arrest).

**Case Study 2: Emotional Decoding for In-Store Advertising.**
A supermarket chain operating in an EU Member State has revamped its in-store advertising system with a smart camera system operated by Echeloon, a company specialising in targeted advertising. Through a camera integrated into a screen displaying advertisements, the system recognises when and for how long a person looks at the screen, their sex (and what it presumes to be their gender), approximate age and worn attire. Furthermore, the system deduces the customer's presumed emotional state (anger, happiness, anxiety, etc.) from their facial expression. The data is then used to personalise advertisements to pre-defined groups of customers and their presumed interests and preferences. Additionally, the system can promote special offers to certain groups of customers or offer specific rebates to an individual customer.

Customers are informed about video surveillance at the entrance of the market where the terms and conditions are posted on signs. However, they contain no reference to the smart camera system. In a press release the chain stated that the system is operated exclusively by Echeloon and the supermarkets have no access to the data. It goes on to state that the system processes only encrypted data and any photos of customers are processed automatically and deleted once the data has been extracted, after approximately 150 milliseconds. Thus, Echeloon assures, no personal data were collected and there was no obligation to inform customers specifically.[2]

[2] For further discussion on legal, regulatory and ethical issues surrounding 'smart' advertising of various kinds see [11].

### 3.2 Group Discussions: Applying the Data Protection Goals to the Case Studies

The group discussions as well as the legal background of the case studies will be discussed in the following. In order to determine whether there is a high risk, the controller, in a first step should refer to examples of high risk processing operations provided by the Article 35(3) GDPR, recitals 71, 75 and 91 as well as the Article 29 Working Party [7]. These include the innovative use of technology, data processing on a large scale as well as publicly accessible areas such as privately-owned shopping centres.

While the case study both fall within these examples and it is thus indicated that a high risk is likely and a DPIA should be carried out, the workshop aimed to enable participants to engage in the evaluation of the risks to the rights of individuals, as it is required in step 2.4 of the DPIA framework.

During the group discussions, the workshop participants sought to apply the Data Protection Goals to the case studies, to analyse the risks to the rights of individuals posed by the processing operations described in the cases. This section summarizes the results of these discussions. Due to the natural flow of the debate, not all protection goals were given equal attention.

The question of whether a processing operation is lawful is paramount in data processing. It is thus assessed as one of the first steps in a DPIA (see step 1.5 of the DPIA framework). While the case studies pose serious questions as to their lawfulness,[3] this was not a focus of the workshop and was therefore not addressed in detail. As the workshop focused on the identification of risks to the rights of individuals, the case studies did not include specific information on legal bases. However, as will be seen below the protection goals are able to identify the risks to the rights of individuals caused by the processing operations lined out in the case studies, which also uncover risks concerning the lawfulness of the processing as the SDM operationalises the legal requirements of the GDPR.

**Case Study 1: Smart Surveillance in Train Stations.**

[3] Participants discussed that while the first case study would have to be based on an express legal basis of national law, the scope of the processing raised serious question of its proportionality. Regarding the second case study participants pointed out that Echeloon's claim that the data processed was not personal was not true, as the duration for which data are processed, and whether or not they are encrypted, is irrelevant to whether the data classify as personal data in the sense of Article 4(1) GDPR. Participants pointed out that the general information at the entrance of the store was not sufficient to obtain informed consent within the meaning of Article 7 GDPR, as it included only general information on video surveillance and not the specific processing operation of emotional decoding. It should be added that just as in the first case study, as the system identifies individuals by use of biometric data, it processes special categories of data according to Article 9 GDPR and thus explicit consent would be required. Beyond these issues, the case further raises issues of price discrimination based on age, gender, race or income (through the analysis of worn attire).

*Data Minimisation.* The first group found that there are several issues concerning data minimisation, which are linked to the extremely broad purpose of the smart surveillance system. It is supposed to identify not only various kinds of offenders, suspects or persons of interests included in a police database, but also any individual in the train station that acts suspiciously and thus allows for the tracking of all passengers frequenting the train station. Furthermore, the system is supposed to alert authorities of dangerous behaviour to prevent harm to individuals. As participants pointed out, these sweeping purposes can already be seen as colliding with the principle of purpose limitation of Article 5(1)(b) GDPR.

Participants also found that the storage of the cameras' raw data in a centralised system for one year violated the principle of data minimisation, as it was not specified why the data would be needed retrospectively, if the person identified did not lead to a match with the police database or act suspiciously or dangerously. Essentially, storage of the raw data beyond the assessment of their identity/behaviour would entail mass-scale data retention on train passengers, the vast majority of whom are neither suspects nor persons of interest. Furthermore, it was questioned, whether the purposes of identifying suspects or persons or interest and dangerous behaviour could not be met through other, less data-invasive means than the proposed smart-camera system.

*Availability.* As to the availability of the data it was discussed that it had to be ensured that the automated algorithm that automatically notifies the pre-defined authorities is revisable and allows for review of its functionality by the controller, e.g. through a logging mechanism. However, the contents of these logs should, with regard to the protection goal of data minimisation, only log data that are necessary to monitor the correct functioning of the system.

*Integrity.* Concerning the integrity of the data processing operation, the first group had general concerns about the properties of the system and the cameras and their safety and security. As the case study did not specify any of these issues an actual system would have to ensure that the entire surveillance system continuously complies with the specifications (including a definition of data flows, concerning access and sharing of the data) and the data processed in it would remain complete or any changes made by employees or external parties could be traced. In this regard the participants emphasised further that, given the amount of passengers frequenting a major train station, even a highly accurate algorithm would produce a significant amount of false positives and false negatives. Hence, it would have to be ensured that these are minimised and the persons operating the surveillance system would be able to adequately interpret these results in order to avoid the risk of false accusations against train passengers. However, the complicated nature of human-machine interactions – especially in the context of hierarchical organisations – exacerbates the risks that false positives or other analytical errors pose to data subjects. Given the complicated and 'inhuman' nature of machine 'thought' [16, 17], staff operating and responding to the system can be presumed not to have more than a rudimentary understanding of how the system reaches its conclusions. Given that they are by definition also likely to hold only low-ranking and possibly insecure positions in their organisa-

tion, they will likely be highly reluctant to question or go against the conclusions drawn by the system: even in the event of the system reaching very questionable conclusions, they will likely have organisational incentives to go along with the machine's conclusions, rather than go against the machine.

*Confidentiality.* The surveillance system entailed multiple risks with regard to the protection goal of confidentiality, the participants of the first group found. Given the broad database, access to the data would have to be defined restrictively and authorized access would have to be logged. This was needed in order to ensure that misuse of the collected data could be prevented or at least be detected and prosecuted. Persons with access to the system would be able to track the daily movements of a vast amount of people. This of course, was not only limited to the controller, who could also be tempted to expand the purposes of the processing even further, but also made the system a high-level target for third party attackers and hackers. A further point of concern was the interface of the system, such as when dangerous or suspicious behaviour is identified and interventions by the station police or security personnel are triggered. Participants stated that it would be a crucial question how much and which data about the individual concerned were made available to the security staff.

Further risks to the rights of the individuals could emanate from the storage location of the data. Participants argued that if, for example, the data were to be stored in a cloud rather than locally, the risk to the confidentiality of the data would be increased even further. Participants again pointed to the risks presented by false positives, false associations, and the potential for bias and subjectivity to infect the analysis. Given the very large number of individuals passing through major train stations, even error rates of less than 1 percent can quickly result in thousands of misidentifications with potentially very serious consequences for the individuals concerned and could subject them to discrimination.

*Unlinkability.* Due to the already overly broad purpose of the surveillance system, the participants focussed especially on unlinkability. The automated matching of individuals with the entire police database was seen as a heavy interference with the rights of individuals. Further, the possibility to identify any individual by matching their photo to the national ID card database was seen as yet another heavy interference with the rights of individuals on a mass-scale. The participants argued that the processed data could easily be used beyond their original purpose in order to discriminate certain groups of people. Due to the raw data of the camera footage being stored, this could also be done retroactively and the data could be combined with data from other sources to track the movements of individuals. Additionally, the data flows and the authorities that can access the data were not sufficiently clear. Lastly, the purpose of the collection could be expanded even further and the system could be linked to other state systems, for instance those of the welfare or health authorities, for instance to monitoring welfare recipients for signs of undeclared employment or other benefit fraud.

*Transparency.* Concerning the protection goal of transparency, it was argued that the train passengers were confronted with the risk of not knowing when, how or why their data was being processed. The individuals would have to be informed of the fact and the amount of surveillance as well as how the data is processed, including whether it is shared with other authorities or private parties, the participants found in their discussion. This had to include the monitoring and/or certification of the algorithm that carries out the biometric recognition and behavioural analysis.

Due to high numbers of individuals concerned they were already subject to a risk of being falsely identified as a suspicious person or as behaving dangerously, especially as these terms were not defined sufficiently. Furthermore, individuals could be identified merely because a person of interest for the police would ask them for the time, as one participant remarked, or the algorithm would identify their behaviour as dangerous. Thus, there was the additional risk of not being able to determine when an individual's behaviour would be registered by the system.

*Intervenability.* Similarly, the individuals faced risks concerning their possibilities of intervention with regard to the surveillance system. The participants argued that the lack of transparency led directly to a risk of the data subjects' not being able to exercise their rights. Furthermore, there was no second instance before the data was shared by the automated system. It was unclear how (and if at all) data subjects who have been identified as suspicious or engaged in dangerous behaviour may challenge a decision, and indeed how they would even find out about such decisions.

**Case Study 2: Emotional Decoding for In-Store Advertising.**

*Data Minimisation.* With regard to the purpose of targeted advertising to customers of a supermarket, the participants of the second group found it questionable whether all of the envisaged categories of data (sex and presumed gender, approximate age, worn attire and emotional state) where strictly necessary, as demanded by the principle of data minimisation. The data collected concern special categories of data according to Article 9 GDPR, as the system uses the biometric data to identify individuals[4] and allows conclusions on categories such as race, ethnicity, religious beliefs (e.g. when wearing a hijab or kippah). Furthermore, the data on the emotional state of customers were derived from the biometric data, could arguably be seen as health data, as Article 9 GDPR includes data relating to mental health (cf. Article 4 (15) GDPR). These broad categories of data, the participants argued, were not necessary to personalise product offers in a supermarket. While the automated deletion of the pictures taken by the system is a step to reduce the amount of data used, the sensitive biometric data is retained indefinitely and therefore the dataset is not reduced to the minimum required to achieve the intended purpose.

*Availability.* The availability of the data here is not an issue, as they are highly available.

[4] Unlike the pictures itself, these data are also stored and further processed.

*Integrity.* Much as in the first case study, the participants of the second group found that concerning the integrity of the data processing operation the properties of the system had to be further defined.

*Confidentiality.* As the data is processed by a processor, the risk of disclosures is higher. Thus, employees of both the controller and the processor could potentially use the biometric data stored in the system for an unspecified period and use them in other processes, such as biometric identification, for identity theft or fraud. Furthermore, other customers or employees could observe the targeted advertisements on the display, which could cause the individual distress, which could, depending on the promoted product, range from mild embarrassment to more serious consequences.

*Unlinkability.* With regard to the storage of the data that is derived from the pictures taken of customers, it was pointed out that the continued storage and further use for other purposes would pose risks to the data subjects, given the nature of the data, which relates to the private life of the individual. For example, if the further processing was aimed at assembling profiles of shopping behaviour – perhaps even drawing on data generated at other stores that use the same camera system – this would amount to tracking of individual preferences.

*Transparency.* In the group discussion transparency was the main issue. The participants argued that the system provided no transparency to data subjects as they were not at all informed about the system. This also extends to the analytical principles governing the system's algorithms: How and on what basis does the system identify certain kinds of behaviour as suspicious or dangerous (including to the individual him or herself)? How reliable is this identification?

The system could also be used to manipulate the emotions of data subjects (e.g., making them unhappy by denying them expected promotions or giving them the 'wrong' ones; making them happy by giving them particular discounts, etc.).

*Intervenability.* As data subjects are not informed of the processing, they would also have no means of intervention in the processing and thus be faced with a negation of their data subject rights.

## 4 Conclusion

The main objective of the workshop was to introduce participants to the DPIA methodology developed by Privacy Forum with a particular focus on the evaluation of risks based on the systematic approach of the SDM, which operationalises the legal requirements of EU data protection law. This was achieved by means of an introductory presentation, and a hands-on exercise in which the workshop participants analysed two data processing operations with regard to the risks they pose to the rights of individuals. As was to be expected both groups found that due to the numerous risks

to the rights of individuals the envisaged processing operations of both case studies could not be carried out.

Beyond the details of the case studies and the particular methodology presented, the workshop discussions yielded insights that are of more general significance for DPIA processes. The discussions among participants confirmed that a multidisciplinary perspective is needed in order to identify and mitigate risks to the rights of individuals in a coherent and holistic manner. The workshop demonstrated that the SDM's data protection goals allow for a structured analysis of risks to the rights of individuals in accordance with the requirements of data protection law. Due to the manifold risks data processing entails such a structured analysis is crucial and at the heart of every DPIA. Nevertheless, the risk analysis in accordance with the GDPR needs further refinement and research. The discussions showed that it can be difficult to discuss risks for rights of individuals, if the legal basis for the processing and the potential risk sources, i.e. attackers, have not been identified beforehand, as stipulated in the DPIA framework. Furthermore, the fine-grained evaluation of the risks to the rights of individuals requires clarification. While recital 75 GDPR refers to the varying likelihood and severity of potential damages, which originated in information security, will have to be adapted in order to allow for the correct application within the fundamental rights framework of the GDPR and in conformity with the requirements of the EU Charter of Fundamental Rights. This future work can then also be integrated in the SDM in order to provide controllers, processors, manufacturers and supervisory authorities with guidelines on how to assess risks to the rights of individuals in practice.

## Acknowledgement

This work is partially funded by the German Ministry of Education and Research within the project 'Forum Privacy and Self-determined Life in the Digital World', https://www.forum-privatheit.de/forum-privatheit-de/index.php.

## References

1. Wright, D., and De Hert, P. (eds.) (2012): Privacy Impact Assessment, Springer, Dordrecht, Heidelberg, London, New York
2. CNIL (Commission Nationale de l'Informatique et des Libertés) (2015): 'Privacy Risk Assessment: Methodology (how to carry out a PIA)', Paris. http://www.cnil.fr/fileadmin/documents/en/CNIL-PIA-1-Methodology.pdf; ICO (Information Commissioner's Office) (2014): 'Conducting privacy impact assessments. Code of practice', UK Information Commissioner's Office, Wilmslow. https://ico.org.uk/media/for-organisations/documents/1595/pia-code-of-practice.pdf
3. European Commission (2011): 'Privacy and Data Protection Impact Assessment Framework for RFID Applications', Brussels. http://cordis.europa.eu/fp7/ict/enet/documents/rfid-pia-framework-final.pdf
4. Friedewald, M. et al. (2016): 'White Paper Datenschutz-Folgenabschätzung', https://www.forum-privatheit.de/forum-privatheit-

de/texte/veroeffentlichungen-des-forums/themenpapiere-white-paper/Forum_Privatheit_White_Paper_Datenschutz-Folgenabschaetzung_2016.pdf

5. Friedewald, M. et al. (2016): 'A Process for Data Protection Impact Assessment under the European General Data Protection Regulation', in: Schiffner, S., et al (eds.) Privacy Technologies and Policy, LNCS 9857, pp. 21-37
6. The Standard Data Protection Model (SDM), V.1.0 EN1 (2017), https://www.datenschutz-mv.de/static/DS/Dateien/Datenschutzmodell/SDM-Methodology_V1_EN1.pdf
7. Article 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'. WP 248, http://ec.europa.eu/newsroom/document.cfm?doc_id=44137
8. Bieker, F. (2018): 'Die Risikoanalyse nach dem neuen EU-Datenschutzrecht und dem Standard-Datenschutzmodell', Datenschutz und Datensicherheit Vol. 42, Issue 1, pp. 27-31.
9. Wright, D., and Friedewald, M. (2013): 'Integrating privacy and ethical impact assessment', Science and Public Policy, Vol. 40, Issue 6, pp. 755-766; Wright, D., Friedewald, M. and Gellert, R. (2015): 'Developing and Testing a Surveillance Impact Assessment Methodology', International Data Privacy Law, Vol. 5, Issue 1, pp. 40-53.
10. Hennen, L. (2012): 'Why do we still need participatory technology assessment?' In: *Poiesis & Praxis* 9.1-2, pp. 27–41. DOI: 10.1007/s10202- 012-0122-5; Slocum, N., Steyaert, S., and Berloznik, R. (2006): *Participatory Methods Toolkit: A practitioner's manual.* Brussels: King Baudouin Foundation
11. Kiesche, E., (2017): 'So funktioniert die Folgenabschätzung', Computer und Arbeit, Vol. 26, Issue 2, pp. 31-36
12. Finn, R. L.; Wadhwa, K. (2014): The ethics of "smart" advertising and regulatory initiatives in the consumer intelligence industry. In: Info Vol. 16, Issue 3, pp. 22-39
13. Somody, B.; Szabó, M.; Székely, I. (2017): 'Moving away from the security-privacy trade-off: The use of the test of proportionality in decision support'. In: Friedewald, M.; Burgess, J. P. et al. (eds.): Surveillance, Privacy and Security: Citizens' Perspectives. London and New York: Routledge (PRIO New Security Series, pp. 155-176.
14. Zielonka, G.: Täglich 450.000 Reisende am Frankfurter Hauptbahnhof, DMM Der Mobilitätsmanager (27 October 2014), http://dmm.travel/news/artikel/lesen/2014/10/taeglich-450000-reisende-am-frankfurter-hauptbahnhof-63731/
15. Probst, T. (2012): ‚Generische Schutzmaßnahmen für Datenschutzziele'. Datenschutz und Datensicherheit Vol. 6, pp. 439-444
16. Burrell, J. (2016): 'How the machine thinks: Understanding opacity in machine learning algorithms', Big Data & Society, pp. 1-12
17. Metz, C.: 'How Google's AI viewed the Move no Human could Understand', Wired (14 March 2016), https://www.wired.com/2016/03/googles-ai-viewed-move-no-human-understand/