European Commission

JRC TECHNICAL REPORTS

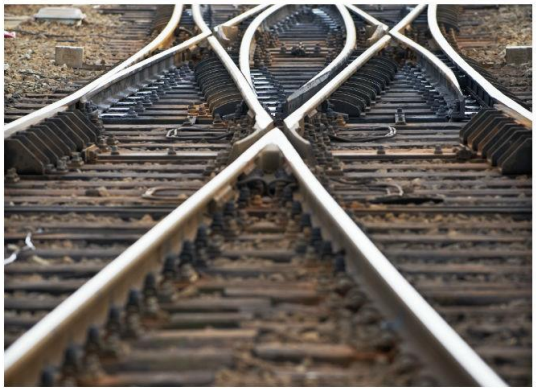
Guidelines for public administrations on location privacy

European Union Location Framework

Francesco Pignatelli
Ray Boguslawski
Leda Bargiotti
Inge Gielis
Bram Verdegem
Paul Smits
Dara Keogh

Version 2
2019






EUR 28202 EN

This publication is a Technical report by the Joint Research Centre, the European Commission's in-house science service. It aims to provide evidence-based scientific support to the European policy-making process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use which might be made of this publication.

Contact information
Name: Francesco Pignatelli
Address: Via E. Fermi, 2749 21023 Ispra (VA), Italy
E-mail: francesco.pignatelli@jrc.ec.europa.eu
Tel.:+39 0332785659

JRC Science Hub
https://ec.europa.eu/jrc

JRC103110

EUR 28202 EN

ISBN 978-92-79-63495-6 (PDF)

ISSN 1831-9424 (online)

doi:10.2791/420310 (online)

How to cite: Leda Bargiotti, Inge Gielis, Bram Verdegem, Pieter Breyne, Francesco Pignatelli, Paul Smits
Ray Boguslawski; Guidelines for public administrations on location privacy; EUR 28202 EN; doi:10.2791/420310

Table of contents

Acknowledgements

This publication is the result of the many discussions, presentations, workshops and analyses carried out in the course of the European Union Location Framework (EULF) and European Location interoperability Solutions for e-Government (ELISE) projects.

The authors wish to thank in particular:

The ISA Working Group for Spatial Information Services and the ISA^2 Working Group for Geospatial Solutions for their interest and support on this important topic.

PWC contributors, Leda Bargiotti, Inge Gielis, Bram Verdegem for their substantive input to prepare the initial version of this guidance document, which gave a location data focus on the broader topic of data privacy and came up with some important recommendations.

Dara Keogh for his contribution from the beginning of 2018, where he has helped the ELISE project in communicating and assessing the topic in many different fora and has discussed practical considerations on preparations for GDPR with a large number of people. He has distilled the learning and ideas generated during this period, to give an increased practical dimension to this second version of the guidelines.

We would also like to thank all those who contributed through interviews, webinars, workshops, on line groups , emails and phone calls into the research that brought new insights, examples and understanding that are now included in the guidelines. While all contributions were equally valuable there are several groups of people whose contributions deserve mention:

a) Those who took part in one-to-one interviews. While we agreed to keep these anonymous, the insights, honesty and reflections were genuinely revealing and greatly helped in our thinking and research;

b) The presenters in two webinars run as part of the outreach programme. These were Efrén Díaz Díaz (Bufete Mas y Calvet), Laila Aslesen (Kartverket) and Yves Schellekens (Agoria). We would like to thank EUROGI for co-ordinating the webinar to their members;

c) The debate and discussion at the Nordic workshop in June 2018, with particular thanks to Ulla Kronborg Mazzoli (SDFE) who lead and organised the workshop and Sik Cambon (SCA) who gave permission for the use the model of linking mapping and anonymisation in the guidelines;

d) The high quality of presentation and engagement at the INSPIRE workshop in September 2018 from Bart De Lathouwer (OGC), Charlotte Fjeldberg (SDFE) and Henri Kujala (Here).

Abstract

Public administrations increasingly use location data to deliver public services, including location-enabled tools, apps for tourists, toll collection services and cadastral web applications. Location data, such as addresses, GPS coordinates or camera images, is key to many public services and can also be linked to all sorts of other data, generating new information that was not available before. Despite the increased consumption of location data, its potential to reveal personal information is often underestimated, especially in comparison to other sensitive data, for instance in the financial and health domains.

Location data not only says where an individual is, it also says who he/she is and what his/her interests and preferences are. Therefore, location data privacy is of paramount importance for public administrations dealing with location data. While location data privacy has many aspects in common with general data protection principles, it also has unique characteristics that require specific consideration.

The goal of this guidance document is therefore twofold: to outline the key obligations that public administrations should comply with when handling personal location data and raising awareness about the importance of location data privacy, highlighting key implications and risks associated with the processing of location data. It does so by guiding the reader through concrete scenarios that public administrations might face when processing personal location data and provides a set of effective and practical recommendations that can help ensure the adequate protection of personal location data.

The guidance has been updated following the introduction of GDPR, taking into account market research in the location industry of the impact of GDPR. The updated document includes new models and concepts as well as using examples throughout to illustrate changes and potential approaches. It is a guide to practitioners and while touching on the key relevant parts of GDPR, it is not a legal document or legal advice.

Keywords: location data privacy, data protection, guidelines, GDPR

1. Introduction

This guidance document addresses public administrations that use or are planning to use location data in their products and services. It provides actionable recommendations to ensure that privacy-related aspects are taken into account when using personal location data.

1.1. Context

This document has been prepared as part of the European Union Location Framework (EULF) action, funded through the European Commission's Interoperability Solutions for the European Public Administrations (ISA) Programme and the follow-on European Interoperability Solutions for e-Government (ELISE) action, funded through the Interoperability Solutions for European Public Administrations, Businesses and Citizens (ISA^2) Programme. The ISA and ISA^2 Programmes support interoperability and sharing and reuse of solutions among European Public Administrations through the creation of, *inter alia*, frameworks, architectures and re-usable components to enable more cost-effective e-Government services and support cross-border applications.

Public administrations increasingly use location data, consciously or unconsciously, to carry out their activities, both for the delivery of public services as well as for internal purposes. However, there may be a lack of awareness of the collection and processing and of location data and the personal data privacy impact as it may be embedded within the overall process, e.g. the multitude of sensors and personal data that is being collected and collated in the management of Smart Cities. Other services include location-based services such as toll systems for vehicles, tourist services or cadastres. Almost every service that is provided contains an element of location, e.g. addresses, GPS coordinates, camera images. Moreover, all sorts of data can be linked to a location, including financial data or health data. In general, location data is closely linked to individuals, which increases the importance of location data privacy.

While location data privacy has many aspects in common with general data protection principles, there are particular characteristics of location that need to be considered:

1. *Identity inference*: location data might not explicitly reveal an individual's identity, but by aggregating disparate data, it is often possible to infer the identity of an individual.
2. *Embedded*: location data is not always the core element of a service but is frequently used implicitly, whether to upgrade the functionality of the service (e.g. location-based service) or for the supplementary use (e.g. direct marketing).
3. *Necessary*: location data is becoming an essential attribute in delivering added value services or products. Individuals expect to benefit from location based services and enjoy a certain level of comfort and automation.
4. *Abundancy*: location data is increasingly used in services and products resulting in large amounts of location data, e.g. location-aware devices (smart phones), digital calendars, personal digital images. As this amount of data rises, it becomes more difficult to manage and control.
5. *Undervalued*: individuals recognise the importance of protecting health or financial data. However, they are not always aware of the value of the data they make available by constantly using their GPS, Wi-Fi and Bluetooth on

their mobile devices. Location data not only says where you are, it says who you are. This situation is perhaps changing as individuals rely increasingly on location-based services and their expectations on data protection develop accordingly (see Section 5 for results of Here and KPMG surveys).

Public administrations need to be prepared and anticipate the risks associated with the use of (personal) location data, keeping these characteristics in mind.

1.2. Target audience

These guidelines target public administrations that use (personal) location data to fulfil their mission. They are conceived particularly to help individuals that have a limited experience with data protection or the specific aspects of location data that may impinge on personal data privacy.

1.3. Scope

The guidelines address the privacy implications of handling location data by public administrations and identify potential risks related to the processing of personal location data. They aim to provide a practical interpretation of legal matters as part of the EU legal framework. The main piece of the EU legal framework that is being considered is the General Data Protection Regulation [2] (GDPR) which entered into application on 25 May 2018. The GDPR replaces the previous data protection directive (officially Directive 95/46/EC1) and lays down the way of data protection in the new digital age.

These guidelines do not however, provide any legal advice nor give an extensive or complete overview of the applicable legal framework. As these guideline want to reach the widest possible audience within public administrations, Member State legislation or sector specific data protection regulation are not taken into account.

A piece of EU sector-specific legislation worth mentioning is the ePrivacy Directive (a.k.a. the "cookie-law"). Drafted in 2002 and updated in 2009, it intends to regulate the telecommunications industry by inter alia protecting the confidentiality of communications, establishing data breach notification and governing the use of cookies, location data and metadata. However, with the introduction of the GDPR, there are some areas of misalignment between the two pieces of legislation (e.g. different time periods for data breach notification or enforcement by data protection authorities rather than telecom regulators). The ePrivacy Directive is currently under revision and a new proposal is expected in 2019. The new ePrivacy Directive is expected to act in unison with GDPR.

Because of the dedicated legal frameworks for law enforcement applications, the processing of (personal) location data in the context of criminal law enforcement and national security is considered out of scope.

1 http://eur-lex.europa.eu/legal-content/EN/LSU/?uri=celex:31995L0046

1.4. Structure of the document

For public administrations it is important to know what the legal obligations and principles are, how these apply to realistic scenarios and what good practices are in the use and management of location data. This document provides guidance on these topics and is structured as follows:

1. Introduction (this section)
2. Executive Summary
3. Definition of location data privacy
4. Key legal obligations arising when handling personal location data;
5. Specific challenges that public administrations face with regard to the processing of personal location data;
6. Recommendations on personal location data;
7. Conclusions of this guidance and summary of the most important elements; and

References

List of abbreviations and definitions

List of figures

List of tables

Annexes

I Three real-life case studies related to the protection of personal location data

II Approaches to data anonymisation.

2. Executive Summary

This is the first update of the guidelines since their initial publication in 2016 pre the introduction of the GDPR legislation. The guidelines have been updated following both the introduction of GDPR and an engagement programme carried out by the ELISE project within the ISA[2] programme focused on location data privacy in both the public and private sectors. This research included interviews, webinars, and presentations and culminated in a workshop on GDPR at the 2018 Inspire Conference in Antwerp, Belgium[2].

The guidelines have been updated and modified to take account of the questions, concerns and potential answers that were raised by this engagement programme with the introduction of GDPR. Some of the questions raised centred around when location data becomes personal data and, if it does, who has responsibility for personal data protection and how or if the responsibilities are shared? Other questions were more direct, for example: should a personal name be included in an organisation's published metadata file? Or if a person's name is in Cadastral file is it personal? During the outreach programme, answers were found that were sometimes similar and sometimes different, the difference often being the context in which the question is asked.

Many of the data protection principles of GDPR existed already before GDPR and many of the principles within GDPR are good data management and governance practices. However, GDPR has extended the definition of personal data to include location as part of personal data and it clearly holds organisations accountable for their action or inaction around personal data that they hold. This created dialogue about future and legacy activities, internal data management and staff training.

The growing dependence of organisations on data to deliver cheaper, better, more targeted and timely services has location at the heart (as all activity happens in a place). Location is a key piece of data infrastructure that in a world of increasing hyper interconnectivity has become a key public concern. Recent surveys from Here and KPMG show concern around personal privacy and low confidence in the enforcement of data protection for location data. This can lead to what researchers refer to as 'a lived reality'. In this, individuals feel that they move between a state of 'Surveillance' to one of control 'Anxiety'. However, a key tenet of GDPR is to build trust in digital services by addressing these concerns and placing the control of their data back to the individual.

The fundamental shift in the understanding of who owns the data changes the emphasis of how and why the data is used and managed. Public bodies frequently rely on legislation to both guide and provide legality to what they do. However, under GDPR each person has a much larger control of their data with rights such as the right to be forgotten, correction of records as well as consent. While these rights are not always what are called 'Absolute Rights', they do create a significant shift in emphasis of data ownership. GDPR move us towards a more Open Data definition of data where it moves from being owned by an organisation to being under the stewardship of an organisation.

GDPR is not happening in a vacuum - the world is more complex and interrelated with public administrations becoming more dependent on private organisations for

[2] https://inspire.ec.europa.eu/conference2018/view-workshops

services, technologies and data. This is particularly clear in the police services where, for example, the tracking of a person by following their mobile phone movement has been instrumental in achieving convictions. A challenge in this context is the increased need for expertise and resources to tackle the issues GDPR raises. The Espresso Report[3] in 2018 highlights the tendency of Smart Cities to push the preparation for GDPR onto private industry. However, while this may work for externally-focused activity and new activity, there are many internal and legacy systems and data silos in public administrations that need to be assessed and aligned with GDPR.

The purpose of this document is to help practitioners in their work with location data under GDPR rules. It is a guide not a rule book and, in all cases, if you are in any doubt you should consult your legal advisors. The guidelines highlight the key points from GDPR and how they apply to location data as well as the policies within organisations. It is intended that the document will be updated with new references and examples, as the application of GDPR in public administrations develops.

The recommendations (Section 6) are:

- Set up a governance structure that enables the development of policies, strategies and guidelines to provide clear direction;
- Set up a data management programme to act on the detail;
- Recognise that data subjects are always the data owners;
- Create trust through transparency with clear, concise and straight forward language;
- Publish a privacy notice;
- Create a focus on privacy which will include security;
- Recognise that protections are only as secure as the weakest link;
- Reduce privacy risks to acceptable levels; and
- Prepare for the worst and you will not be caught out.

In summary, GDPR has introduced a new layer of complexity that provides the opportunity to build better data practices and trust in public digital services with all the associated benefits. For example; DG Just estimate that GDPR will result in €2.3 billion of savings in reduced administration costs for organisations operating across two or more European countries. It means putting in good data management and structures in place, a change in perception and a move from data ownership to data stewardship and the inclusion of location data at the core of all the changes.

[3] http://espresso-project.eu

3. What is location data privacy?

Location data privacy has no clear-cut legal definition. For this study, we derive a definition from the definitions of personal data and location data both of which are defined in the European Union legal framework.

Personal data

Definitions of personal data are available in various sources. Throughout this document, we will use the definition provided by the General Data Protection Regulation (GDPR) [2] adopted in May 2016. Article 4 of the GDPR defines personal data as follows:

''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;

Location data

In the context of this document, a broad view of location data has been taken to encompass as many scenarios as possible. The guidelines include any data with an implicit or explicit geographic or geospatial reference, ranging from address data to radio signal-based triangulation or IP address location, including data published under the INSPIRE Directive [4].

Personal location data

For the purpose of this document, personal location data is any location data directly or indirectly linked to a living individual or that can be directly or indirectly used to identify a living individual. GDPR only applies to living people or 'data subjects' it does not apply to the dead or to organisations. In other words, organisations do not and cannot have personal information rights under GDPR. However, the personal information of the employees of an organisation is protected by GDPR.

Figure 1: *Relationship between personal data and location data*

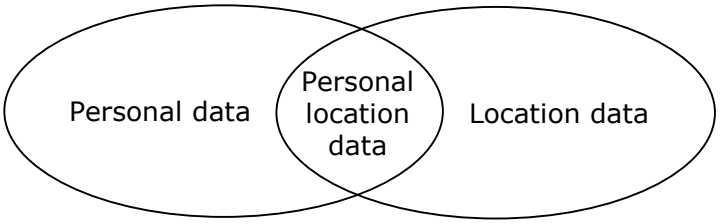


Figure 1 above shows the three areas of consideration; Personal data, Location data and where they intersect i.e. Personal Location Data.

In taking this broad approach it is important to draw a distinction between data that has location information and data that could have personal location information when combined with other data. It may be that in the combining of datasets the location data becomes personal information. For example; under the INSPIRE themes, Elevation may not contain any personal information. However, if you link to Land Use, Natural Risk Zones, Sea Regions and Geography Names you may quickly narrow down

the data to a personal level and identify where somebody lives or works. See Table 2 below for some more examples.

Table 1: *Examples of personal location data*

Example of location data	Example of personal data	Personal location data
GPS coordinates of the location of a smart phone	Telephone subscription account information linked to the smart phone	By combining the two data sources, the location of the individual can be identified.
Public IP address	Internet subscription account information	
Cadastral information about a realty	Realty owner information	
Traffic camera footage on a specific location	Licence plate owner information	

Location data privacy

Location data privacy is the individual's right not to be subjected to unauthorised collection, aggregation, processing and distribution (including selling) of his location data. It is the right to be protected by the ability to conceal information of whereabouts, which can be derived from personal location data.

While many of these rights previously existed, they have been extended to include location data as personal data and they have pushed greater responsibility and accountability of the management of personal data on to organisations. As the table below shows, GDPR is the most comprehensive piece of data protection and it has been described as the gold standard of data protection.

Table 3 GDPR and Previous Data Privacy Provisions (source: ESPRESSO Project)

	Privacy by design	Certifications	Security	Data Ownership
Data Protection Directive (1995)			✓	✓
ePrivacy Directive (2002)		✓	✓	
GDPR (May 2018)	✓	✓	✓	✓

4. Legal obligations when processing personal (location) data

The protection of personal data is a fundamental right. New technologies introduce new privacy risks and a more privacy-aware and assertive society requires fully fledged control measures to protect private life. Policy makers and standardisation organisations are responding to this by updating data protection legislation and guidelines.

The right to the protection of personal data however is not an absolute right. It should always be considered in relation to other fundamental rights such as freedom of expression and information, or freedom to conduct a business. Specific to public administrations is the right to access public sector information such as official documents, which may contain personal data.

While technically and legally the protection of personal data is not an absolute right the principles and spirit of GDPR put an emphasis on protecting the individual's privacy and in placing the individual back in control of their data. This position should always be considered in relation to other fundamental rights such as freedom of expression and information, or freedom to conduct a business. Specific to public administrations is the right to access public sector information such as official documents, which may contain personal data.

This chapter consolidates the main privacy obligations that public administrations should comply with when processing personal location data. Each obligation is structured in three parts: (1) **explanation**, (2) **example** of how the obligation applies to location data and (3) **references** to specific data protection regulations and guidelines. The following example is used throughout to illustrate the concrete application of each principle with regard to location data privacy:

A public Tourist Information Office provides visitors with a paying smartphone app for guided tours through the city. The app uses GPS coordinates to define the visitor's location and indicates on a map the route to take. The app has other functionalities as well, such as suggesting where to have a drink or take a meal, to rate restaurants and pubs and to evaluate tourist attractions. All the information provided by the users of the app is hosted on a central system, owned by the responsible public administration for tourism.

In working through the text it has been assumed that the resources, budgets and skills are available to meet the various needs of any project. It is also assumed that technology, private organisations, and workable solutions are available.

4.1. Appoint a responsible individual for data protection

Appoint a responsible individual for data protection within your organisation, to supervise the management of personal location data and provide the necessary level of transparency within the organisation and towards data subjects. The responsibilities should encompass a number of tasks from strategy to execution, including but not limited to: counselling, defining policies, monitoring compliance, raising awareness and training staff, executing privacy risk assessments or communicating with supervisory authorities and data subjects.

According to the General Data Protection Regulation [2] each public administration shall appoint a Data Protection Officer (DPO).

Example

The Tourist Information Office could appoint its own DPO. An alternative option could be to appoint one DPO for all Tourist Information Offices in the region or the whole country, depending on the total number of Tourist Information Offices. Additional national legislation might also give more guidance on this. The role of the DPO does not have to be a full time function; however the time and effort the DPO spends on data privacy should be in line with the extent of the app usage or other personal data processing activities performed by the Tourist Information Office. The workload of the DPO could be impacted by e.g. the number of app users, the amount of processed personal data, the number of complaints or the pace at which new functionalities are introduced (and need to be assessed). The tasks of the DPO can be combined with other tasks in the organisation, as long as there is no conflict of interest.

Related regulation or guidelines

- GDPR [2], article 37: Designation of the data protection officer
- GDPR [2], article 38: Position of the data protection officer
- GDPR [2], article 39: Tasks of the data protection officer
- Regulation 45/2001 [7], article 24: Appointment and tasks of the Data Protection Officer
- Regulation 45/2001 [7], article 25: Notification to the Data Protection Officer
- OECD Privacy Framework [5], Part II, Accountability Principle

4.2. Ensure lawful processing of personal location data

The processing of personal location data has to be lawful and fair (amongst other things, individuals may not be deceived or misled) and has to be transparent in relation to the data subject.

In particular, public authorities and bodies can lawfully process personal (location) data if they have a legal basis, i.e. if the processing is necessary for the performance of a contract of which the data subject is party, if the processing is necessary to comply with their legal obligation, if it is necessary to protect the vital interests of a natural person or if it is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller, where and if required a data subject has given consent for the specific purpose.

This means that public administrations do not have to ask for the consent of the data subject if there is a legal basis or legislative measure that regulates the processing of personal location data for performing a specific task. However, this legal base does not imply that the other applicable privacy principles (see the other obligations in this section) become irrelevant. Lawful processing is only one of many mandatory requirements. It also should be noted that while data maybe collected under legislation for one purpose does not mean it can be used for another purpose.

A question that arose during the research and after the introduction of GDPR, was whether a personal name should be included in metadata files that are required under for example the INSPIRE Directive. The answer is twofold. Firstly, it is good practice and there should be a contact point in the metadata. However, it may be better to address it to a universal mailbox of the type 'info@xxxxx' This is to ensure someone picks up the question as an individual may leave an organisation and their email

address becomes redundant, resulting in queries potentially being lost. As long as the query is picked up and assigned to the correct role it may be a better option. Secondly, if a personal name is included under GDPR the email is personal and it should only be used for the purpose for which it was given, namely for queries re metadata or other technical queries on the data. The email address cannot be used and harvested for any other purpose and the person can object and have it removed from any irrelevant circulation list where it is included.

Example

In order for the Tourist Information Office to process location data lawfully, there must be a legal basis or legislative measure justifying the processing activities. It might be legal for the Tourist Information Office to store a limited set of personal location data, but most likely there is no legal context or public interest to process numerous detailed personal location data that is collected through the app. In the absence of a legal basis or legislative measure, the Tourist Information Office can however process the data lawfully only if (1) it asks users for explicit consent or (2) it completely anonymises all collected data so it cannot be tracked back to an identifiable individual.

Related regulation or guidelines

- GDPR [2], article 5: Principles for personal data processing
- GDPR [2], article 6: Lawfulness of processing
- Regulation 45/2001 [7], article 4: Data quality
- Regulation 45/2001 [6], article 5: Lawfulness of processing
- Data Protection Directive [7], article 7
- OECD Privacy Framework [5], Part II, Collection Limitation Principle
- OECD Privacy Framework [5], Part II, Use Limitation Principle

4.3. Apply data protection by design and default

This is about building data protection into the design and concept phase of any project. It is about enshrining personal data tests at the heart of the product/service development process so that key project questions become; 'does this affect privacy?' and if it does 'how do we minimise the use of personal data?', as well as a series of associated questions such as how long do we keep the data and how do we let the user know what their data is to be used for and why.

Within GDPR there are two distinct users of data and these are referred to as data controllers and data processors. Without going through all the detail, a quick and short way to think of this is that processors do as they are told by the controllers. Both have responsibilities under GDPR. The controller has the greater responsibility as they determine the purpose and the means of processing the data or in other words the how and the why of data processing. The third option catered for under GDPR is that organisations as considered to be controllers of the data are therefore treated under GDPR as joint controllers. An interesting question that arises is when do you stop being a controller? For example, if data you provide has no personal information but is then combined with another or multiple data sets, who has responsibility? The general opinion from research for this document is that the organisation that does the additional processing and combining of the files becomes the Controller of the

combined files. However, even if an organisations anonymised data is used by another organisation, GDPR puts a responsibility on the issuer of the original dataset to continue to ensure that the anonymisation still holds. This may entail the issuer of the data having to review the level and techniques used for anonymisation (see Annex II).

To ensure that data protection by design and by default is an overarching principle, it is important to include privacy risks when drafting the business case and to follow up on them during the progress of the project. When designing a new solution or service, technical (e.g. use of pseudonymisation or anonymisation software) and/or organisational (e.g. process for complaints handling) privacy controls could be added. Privacy by design should be a default mind-set which is established within an organisation and project team.

Example

The Tourist Information Office should take data protection into account from the moment the idea of creating an app for guided tours is conceived. This can be done by asking some basic questions: *What is the goal of the app? What personal data do we actually need? To what privacy risks will users be exposed by using our app? What safeguards do we need to put in place to protect app users' personal data?* By asking these questions upfront, the Tourist Information Office avoids major changes to its app during the development phase, or even worse, when already in use.

Related regulation or guidelines

- GDPR [2], article 23: Data protection by design and by default
- Information and Privacy Commissioner of Ontario – Privacy by Design (PbD) [8]
- ENISA - Privacy and data protection by design [9], chapter 3: Privacy Design Strategies

4.4. Apply data minimisation

Only adequate and relevant location data can be collected and processed. The collection must be limited to what is strictly necessary for the purpose of which data was collected in the first place. The challenge consists in establishing the right level of specificity and granularity of location data needed for the service or product.

Redundant data should not be collected and aging data no longer be used and should be deleted in line with the data retention policy of the organisation. Unfortunately, there is no mathematical formula to calculate the retention period. When defining the retention period, a balance should be made between the protection of individual's privacy and the public administration's needs for which the personal data was collected.

It is a fundamental requirement that personal location data collected for one purpose cannot be retained once that initial purpose has ceased nor can it be used directly, indirectly or incorporated into another solution without either having a legislative, contractual or the users consent to do so.

Example

The Tourist Information Office wants to improve its app's functionality by suggesting tourist routes based on the typical interest of a tourist age category. Therefore, they want to ask users for their birthday to see what touristic attractions are more popular

with a certain age category. However, for this purpose it would be sufficient if users indicate in what age category (e.g. ages 18–21; 22-28; 29-36; …) they are situated instead of asking for their day of birth.

As the Tourist Information Office is collecting personal (location) data, it must also communicate to its app users for how long all the data will be retained. There can be different retention periods for different types of data or some basic data could be retained for a longer time, even if the app user deleted his profile. If the Tourist Information Office anonymises the data it retains when the initial purpose has ceased, there are no limitations as it is no longer classified as personal data. This assumes that the data cannot be reversed back to personal data and in effect it is no longer personal data.

Related regulation or guidelines

- GDPR [2], article 5: Principles for personal data processing
- GDPR [2], article 17: Right to erasure / Right to be forgotten
- Data Protection Directive [7], article 6 (c)
- OECD Privacy Framework [5], Part II, Collection Limitation Principle
- OECD Privacy Framework [5], Part II, Purpose Specification Principle

4.5. Perform periodic privacy risk assessments

Unfortunately, risks are not static and they change and develop over time. This can be due to technological, product/services and/or legal changes. So if a system receives an upgrade, part of the upgrade should include a review of privacy which under GDPR is called Data Privacy Impact Assessment. This is to ensure an accurate level of data protection towards data subjects (you and I). Public administrations should assess the risks they expose data subjects to when processing their location data. This is not a one-off activity. As risks evolve, the likelihood and impact of these risks should be re-assessed regularly.

Next to the risks data subjects are exposed to, public administrations expose themselves to risks as well when processing personal location data. The risk of e.g. non-compliance, data leakage, insufficient or ineffective security controls should be assessed as well.

This risk assessment usually takes the form of a Data Protection Impact Assessment (DPIA) or Privacy Impact Assessment (PIA).

Example

Throughout the life cycle of the app, the Tourist Information Office should identify and assess privacy risks, at least yearly or whenever the app undergoes a significant change or when a new venerability or security risk is identified. New functionalities might introduce new privacy risks, threats and vulnerabilities may evolve or new ones arise, and also, people's privacy risk appetite changes over time.

An example of a new functionality for the app could be the integration with social media platforms (e.g. Facebook or Instagram) to share information with your friends (or the public), such as for example; the tourist route you have been travelling or the next stop on your tourist route you are heading to. The implementation of this functionality should be subject to a privacy risk assessment. As probably not every

individual will make use of this new functionality, the app user should have the option to opt-in or opt-out.

Under GDPR, it is necessary to always ask the user's consent and enable them to adjust this consent to access any other app or service on their smart phone. For example, in order for the tourist app to work, it will need access to the user's location information, and permission must be sought to do this.

A simple message along the following lines would work:

'In order for this App to work we need access to your location so we can show you the nearest tourist attractions to you and enable you to plan your touring activities'.

Related regulation or guidelines

- GDPR [2], article 35: Data protection impact assessment
- Location Data Privacy Guidelines [10], part 4: Location Data Privacy Risk & Transparency Assessment

4.6. Secure data processing activities

Irrespective of the lawfulness of the processing of personal location data, processing activities should be secured adequately. As described above, the mandatory privacy risk assessment identifies privacy risks. These risks must be mitigated through the use of technical and/or organisational security controls. Commonly mentioned security controls are encryption or pseudonymisation [Annex II], but all well-established security principles such as 'need-to-know' (*i.e. allowing access to information or knowledge only if required to perform an assigned task*) or 'layered security' (*i.e. a defensive security strategy featuring multiple layers that are designed to slow down a security attack*) contribute to the overall level of security. Important to know is that the overall level of security of a solution is only as strong as the weakest link. This implies every component of a solution, whether central systems or remote devices, should be secured adequately.

There are many security control frameworks that a DPO can refer to. Some of these focus on protecting personal data, such as ISO 27018[4]. Other more general frameworks, such as the ISO 27000 family of standards[5], ISF Standard of Good Practices[6], NIST[7] or SANS[8] publications, are applicable as well.

Example

The Tourist Information Office should safeguard data at three levels to secure its service: data residing on the smartphone, data in transit when transferring it to the central system, and data residing on the central system. It seems appropriate to apply typical safeguards such as developing the app according to secure coding principles, encrypting the data transferred to the central system or limiting access to the central system on a need-to-know basis. These are only a subset of adequate security controls, but there are many more as referred to in the paragraph above.

[4] ISO/IEC 27018:2014: Information technology -- Security techniques -- Code of practice for protection of personally identifiable information (PII) in public clouds acting as PII processors
[5] http://www.iso.org/iso/iso27001
[6] Information Security Forum (ISF) - https://www.securityforum.org/tool/the-isf-standardrmation-security/
[7] National Institute of Standards and Technology (NIST) http://csrc.nist.gov/publications/PubsSPs.html
[8] https://www.sans.org

Related regulation or guidelines

- GDPR [2], article 32: Security of processing
- Regulation 45/2001 [7], article 22: Security of processing
- Data Protection Directive [7], article 16: Confidentiality of processing
- Data Protection Directive [7], article 17: Security of processing
- OECD Privacy Framework [5], Part II, Security Safeguards Principle
- ENISA - Privacy and Data Protection by Design [9], chapter 4: Privacy Techniques

4.7. Comply with data subjects' rights

Data subjects remain owners of their personal location data (elaborated on in 6.3). This means they can invoke numerous rights related to that ownership. This section briefly describes the rights of data subjects that data controllers (i.e. public administrations processing personal location data) should respect. These rights are described in full details in the respective articles of the GDPR [2], referred to in the remainder of this section. In general, data subjects have the right to:

- *Access their data*: Obtain information regarding the processing of their data upon request. This includes, amongst others, the data being processed itself, where data is processed and why it is processed, the recipients to whom the data have been or will be disclosed, or for how long the data will be stored. (GDPR [2], article 15, *right of access*). This should be clearly explained on the Privacy Statement on the App. Other information that should be included is what personal data is accessed, how it is used, who it is shared with and why it is used or shared and how long is it kept, why and for what purpose. To enable the user to contact, obtain clarification and/or correct data it is important to have a 'Contact Us' section/button on the App. This should be easy to find and use.
- *Correct their data*: Have inaccurate or incomplete data relating to them rectified. (GDPR [2], article 16, *right to rectification*). This is why the 'contact', button mentioned in point 1 above is important.
- *Erase their data*: Have the data erased when the collected data is no longer necessary, the data subject withdraws consent, the data subject objects to the processing, or the data has been processed unlawfully or for compliance reasons. (GDPR [2], article 17, *right to erasure / right to be forgotten*). This is again why the 'contact', button mentioned above is important.
- *Withdraw their consent*: Withdraw consent for processing data at any time. The withdrawal of consent should be as easy as giving the consent. (GDPR [2], article 7, *right to withdraw consent*). In the App example, a user can turn off their location so it is not shared either when not using the App or in all cases. If the location is turned off when they are not using the App, the App must be designed to do this and only access the user's location when the App is in use. If the user turns off location completely or does not allow the App to access their location data, it may be that the App can launch but provide no functionality. This needs to explained in the Privacy statement on the App,

as a message on the App when the user's location is turned off and when first asking permission to access the user's location.

- *Restrict the processing of their data*: Restrict the data processing in certain circumstances such as unlawful processing or withdrawal of consent. When the processing has been restricted, the personal location data can only be stored but not processed until the restriction is lifted. (GDPR [2], article 18, *right to restriction of processing*). Please note in the example of the App the user's location should only be stored while the user is using the App and deleted as soon as their session is completed on the App, **unless** it is made **very clear** when signing to use the App that the users location information will be kept, for what purpose, and for how long. However, in the case of our tourist App there is probably no need to retain this information at all. What is likely to be more useful is an anonymised record of places looked up and visited on a day to enable better configuration of the App, taking account of what is interesting tourists at different times of the year. This may be as simple as a statement along the following lines: 'we use your location information to improve the performance and enjoyment of App.'
- *Data portability:* If the processing is carried out by automated means and if it is based on consent or a contract, the data subject has the right to receive the data and transmit this data to another data controller. (GDPR [2], article 20, *right to data portability*). This may be difficult to do in practice as there are likely to be multiple standards, languages and data structures involved which may work against this being possible for some or all of the data involved.
- *Make an objection*: In certain situations (e.g. direct marketing, statistical purposes) data subjects have the right to object to the processing of their personal location data. (GDPR [2], article 21, *right to object*). In our App example the user may switch off their location information, delete the App and or request an access request to see what information is held about them.
- *Not to be subject to a decision based solely on automated processing*: The right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her. (GDPR [2], article 22, *right not to be subject to a decision based solely on automated processing*)
- *Lodge a complaint*: The right to lodge a complaint with a supervisory authority if the data subject considers that the processing of personal information infringes the GDPR. (GDPR [2], article 77, *right not to lodge a complaint with a supervisory authority).* This again points to the importance of providing the ability to easily contact the organisation.
- *Effective judicial remedy against a supervisory authority*: The right to an effective judicial remedy against a legally binding decision of a supervisory authority. (GDPR [2], article 78, *right to a judicial complaint with a supervisory authority*)
- *Effective judicial remedy against a controller or processor*: The right to an effective judicial remedy if data subjects consider that their rights have been infringed. (GDPR [2], article 79, *right to a judicial complaint with a controller or processor*)

- *Compensation and liability*: A data subject who has suffered material or immaterial damage shall have the right to receive compensation from the data controller or data processor for the damage suffered. (GDPR [2], article 82, *right to compensation and liability*)

To respect these rights, public administrations should identify the data subject's rights applicable for the specific case and implement the necessary internal processes and procedures. This is achieved by having terms and conditions for the App, a privacy statement and contact section, and a data access request form and process. In the App example, a clear unambiguous statement requesting access to the user's location (see first point above), providing the user with the ability to control this access and easy ways to contact the organisation are practical ways in which the user's rights are respected.

While the right of privacy is not an absolute right as data subjects cannot invoke the right to object to legal tasks carried out by public administrations, it is still best practice and in line with GDPR to outline why the data is being collected, what it is going to be used for, how long it will be retained, who it will be shared with and how it will benefit the user. This is the type of information that will be contained in the organisations Privacy Policy and help build the transparency and trust that is a key driver behind GDPR.

Example

The users of the tourist app have several rights when it comes to their data. In order to respect these rights, the Tourist Information Office should provide app users with the option to e.g.:

- limit the processing of personal location data by limiting the required data,
- delete their restaurant rating or comments to a tourist attraction when they withdraw their consent,
- allow them to update/rectify their personal data stored by the app, or
- access all their personal location data collected through the app.

Related regulation or guidelines

- Regulation 45/2001 [7], article 13: Right of access
- Regulation 45/2001 [7], article 14: Rectification
- Regulation 45/2001 [7], article 15: Blocking
- Regulation 45/2001 [7], article 16: Erasure
- Regulation 45/2001 [7], article 18: The data subject's right to object
- Data Protection Directive [7], article 12: Right of access
- Data Protection Directive [7], article 14: The data subject's right to object
- OECD Privacy Framework [5], Part II, Data Quality Principle
- OECD Privacy Framework [5], Part II, Individual Participation Principle

Note: references to the GDPR [2] are made in the text itself.

4.8. Notify data breaches to data subjects and relevant bodies

Supervisory authorities and data subjects expect to get notified about data breaches. Unless a data breach is unlikely to result in a risk to the rights and freedoms of the affected data subjects (e.g. if the leaked data would be made unintelligible by the use of e.g. encryption), the competent supervisory authority should be informed.

Furthermore, if the data breach would likely result in a *high* risk, the affected data subjects should be informed personally and without undue delay.

The notification should describe the details of the data breach, the control measures already taken, and recommendations for the effected data subjects to control damage. All communication towards data subjects should be transparent and in clear and plain language.

Public administrations might process personal data of numerous data subjects, which makes a personal notification to each data subject practically infeasible. Therefore, a public communication – if effective – is considered to be sufficient.

Example

If the Tourist Information Office was to be hacked and all data of their users were compromised, the Tourist Information Office should first inform the supervisory authority of the data breach. Afterwards and in consultation with the supervisory authority, it should inform its users without undue delay. It could use the application to push a message to all users, or use other communication channels to inform its users that their data is compromised.

Communication in case of data leakage should not be a one-off. It is advisable to provide the affected users and the supervisory authority with regular updates on the progress of controlling the data breach and the measures taken to avoid future data breaches.

Related regulation or guidelines

- GDPR [2], article 33: Notification of a personal data breach to the supervisory authority
- GDPR [2], article 34: Communication of a personal data breach to the data subject
- Data Protection Directive [7], article 10 and 11

5. Using location data: scenarios, challenges and risks

This chapter provides a brief overview of the use of personal location data in the public sector with a focus on the use of this data. The general use is described below, followed by four concrete examples of personal location data processing:

- Location-aware browsing: use of geolocation data
- Electronic eID: use of address data
- Use location data for business intelligence or statistical purposes
- Working with private third parties: exchanging personal location data

For every example, related challenges and risks are identified and reference is made to the privacy principles from section 4 that are linked directly to the example.

In general, we identify three main scenarios where a public administration is involved in the processing of personal location data.

- **Scenario A:** intra-administration - a public administration processes personal location data when providing services in the context of its public task(s) and keeps the data within its organisation. These services can be emergency services, regular public services [15] or law enforcement services.
- **Scenario B:** administration-to-administration - public administration X processes personal location data and exchanges this data with public administration Y.
- **Scenario C:** administration-to-business - public administration X processes personal location data and exchanges this data with private organisation Z.

Independently from the scenarios and examples provided, it is important to highlight the potential added value location-based services might bring. As recent surveys from both KPMG and Here show individuals have a high level of concern about their location data and their privacy. The figures show the value individuals put on their privacy when it comes to location data.

In the KPMG, 2016 worldwide survey[9], of those surveyed:

- 84% are not willing to share their physical location data; and
- 86% are not willing to share their address data.

In the Here, 2017 worldwide survey[10], of those surveyed:

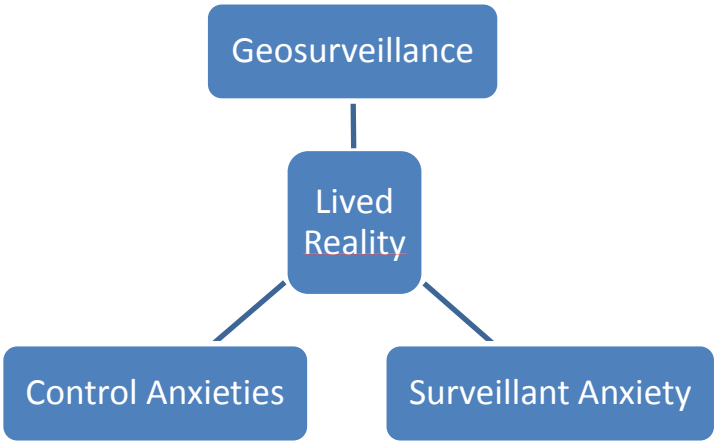
- 75% feel stressed about sharing location information;
- 84% do not trust that the law will protect them against misuse of their location data;
- 71% are willing to share location data with a map or navigation service; and
- 66-68% are more willing to share data if they are clear about how it is being used.

[9] https://home.kpmg.com/sg/en/home/media/press-releases/2016/11/companies-that-fail-to-see-privacy-as-a-business-priority-risk-crossing-the-creepy-line.html
[10] https://www.here.com/en/node/40306

Both surveys show that a sizeable proportion of individuals are wary of sharing their location data. The Here survey also points to the importance of explaining to the user the advantage of sharing their location data and the greater willingness to share location data if users inherently understand the advantage to them.

This moves to one of the core objectives of GDPR which is to build trust with individuals so that they will grow and develop their use of digital services across all branches of government and public services. In reviewing the academic literature on this topic and particularly in relation to Big Data and location the following framework of data subjects' experiences are outlined.

Figure 2: *Location, Big Data and Individuals' Response [11], [12], [13]*



As Figure 2 shows, researchers describe individuals as living in a world of what is called a 'lived reality'. We can feel we are under surveillance which can create anxiety as well as a feeling of loss of control. The researchers indicate that if individuals can access some controls over their data then they are happier.

This experience of having control over your data is a mainstay of GDPR and one that is hoped will build trust in the digital economy. A key way in which organisations can act positively act is by being upfront and transparent about the use of the data. This is supported in the survey results from Here shown above.

Going back to our App example, it should clearly state something like:

'We need access to your location information so that we can show the nearest tourist attractions to you and we will only use your data for this purpose'

This statement is simple, clear, precise, easy to understand and it makes sense to the user to want to share their location as they understand the benefits and purpose.

Some individuals are happy to share their location if there are associated advantages; others are more sensitive to their privacy and might only use a service if their privacy is fully guaranteed. As a service provider, it is important to be able to support both.

5.1. Location-aware browsing: use of geolocation data

5.1.1. *Context*

A public administration offers a service to its citizens through an online portal. To provide a more personalised service, the public administration collects geolocation

data of its citizens' mobile device or computer (e.g. IP addresses or GPS coordinates). This example could easily be expanded to include the multiple sensors in place to enable Smart City actions or the array of sensors and locational information exchange that must occur for autonomous vehicles to function. In both cases personal location data is being accessed, processed and exchanged. In this example we will however, focus on a more generic example of an online portal, the principles of which can be applied to many examples.

5.1.2. Challenges and risks: protection of location data

Citizens are exposed to risks by sharing their location and should always be wary of inappropriate use. A public administration is facing the challenge of protecting the location data that a citizen shares with the public administration. In the case of location-aware browsing geolocation data can, in most cases, only be collected if citizens give their explicit consent. Moreover, public administrations are recommended to provide the option to citizens to choose the level of detail they want to share.

Next to that, a citizen should have the option to withdraw consent. Withdrawing consent implies that a public administration must delete all gathered personal location data of a particular citizen, unless legal grounds for the processing remain.

A specific scenario where geolocation data is used is to block access to (for example copyrighted) content made available through online portals. Based on a visitor's IP address the country from where the access request originates, can be retrieved and blocked if required. Although this action identifies the location (country), a visitor is not traced back to an identified individual. Therefore, no privacy issues arise if content blocking is setup correctly.

5.1.3. Privacy principles

For this specific example the immediately applicable privacy principles are:

- Achieve lawful processing of personal location data: ask for the citizen's explicit consent (see 4.2)
- Apply data minimisation: reflect on the level of detail of location data your provided service requires. Besides that, provide the citizen with the option to choose the level of detailed location data he wants to share (see4.4)
- Comply with the data subject's rights to withdraw their consent or to be forgotten (see 4.7).

5.2. Electronic eID: use of address data

5.2.1. Context

A public administration collects and processes citizens' address data provided via the national electronic eID to fulfil a public task in the context of a legal obligation. This data can be collected in different ways, for example either through an online tool or an application form. In doing this, it should be made clear to the citizen what is happening and why it is happening, including the specific purpose for which the personal data will be used, rather than just quoting reliance on legislation. Remember, the goal is to build trust and grow the use of digital services.

So for example, on using the service we might have a message that says:

'We are required by statue 1.1.2 of Regulation 1.3.5 under clause 3.6a and 4.8b to hold and process the Address Information as defined under Regulation 5'.

This kind of notice is not clear or easy to understand or confidence/trust building. Perhaps the following would be more in line with the motivation of GDPR to build trust and confidence:

'To ensure we give you best service and right information it is essential we use and retain your address information. There are a number of pieces of legislation about this which you can read here'

However, over time the public administration may create a new service and want to promote it to its citizens to point out the opportunities this new service introduces. To do this, the public administration may want to send out personalised information brochures to its citizens whose address data is already available through the collection in the context of its public task. This type of data use can be problematic under GDPR as it may involve the use or re-use of data collected for one purpose to be used for another purpose. Unfortunately, under GDPR this is probably not going to be possible as personal data held and gathered for one purpose cannot then be re-used for another purpose.

As address data is probably the most frequently requested and stored personal (location) data by public administrations, this scenario is closely linked to the once-only principle. This principle implies that citizens should not be asked for the same data more than once, with the objective to reduce the administrative burden for the citizens. However, this does not imply that address information can be shared freely between public administrations and reused for whatever purpose. As described in 4.2, there should always be a lawful base for processing personal (location) data. If a public administration wants to reuse address data, originally collected by another public administration for a specific purpose, the second public administration needs to ascertain whether it can use this piece of data.

5.2.2. Challenges and risks: use of location data for purposes other than the one for which they were collected in the first place

The address data was originally collected lawfully in the context of a public task assigned to the public administration (i.e. for the national electronic identity cards). However, the use of the address data for promoting the new service does not fall under the public task assigned to the public administration. If the public administration wants to use the address data of its citizens in another context, the public administration should explicitly ask for consent.

5.2.3. Privacy principles

For this specific example the key privacy principles are:

- Achieve lawful processing of personal location data: ask for a citizen's explicit consent before using his address data (see 4.1).
- Comply with the data subject's rights to withdraw their consent (see 4.7).

5.3. Use of personal location data for business intelligence or statistical purposes

5.3.1. Context

A public administration has been collecting personal location data for several years in the context of its public task. To improve its service towards the citizens, the public administration decides to subject all collected personal location data to business intelligence processing. The IT department of the public administration has the experience and the tools to provide the necessary support. Moreover, external experts and the software supplier will be involved in this process to achieve high-quality results.

5.3.2. Challenges and risks: Re-identification of anonymised or pseudo-anonymised personal location data

Aside from issues of different data formats and standards, the data issues that have to be faced are: For what purpose was the data gathered? Under what permissions was it gathered? And how long has it been retained? If this information cannot be found, the data may not be able to be used. However, let us assume that the data was gathered under legislation and that the only data that is needed is data that is within any allowable timeframes under any relevant legislation and the organisations data retention policy.

If the data can be used, the next potential risk in this scenario is unlawful disclosure of personal location data to actions by external consultants or suppliers. Therefore, it is important that the public administration takes the necessary measures to protect the citizens' personal (location) data. To do this it should:

- Verify to what extent other parties need the full data set or if a limited data set would be sufficient;
- Decide if it is appropriate to apply data anonymisation on the personal location data;
- Ensure that the correct technical controls and contractual agreements for external staff on non-disclosure and acceptable use are put in place.

A lower risk, but one that needs to be considered, is the risk of re-identifying individuals even if the data is anonymised. Typically, this could happen when a dataset is minimal or granular. For instance, the inclusion of geographic details in a dataset makes it much easier to re-identify users that live in small geographic areas. If a public administration publishes research results on farm subsidies by area and there is only one farm in a particular area, the farm owner can easily be identified.

There are different actions that public administrations can take to limit such risk, which all essentially mean reducing the accuracy of the data and taking away some of the information and value of the data. For example: delete records of small areas, remove from the disclosed dataset some of the non-geographic variables, reduce the precision of geographical areas or aggregate the small geographic areas into larger ones.

Each option has certain disadvantages with regard to the richness of the dataset, therefore when making a decision, public administrations should make a careful assessment.

5.3.3. Privacy principles

For this specific example the key privacy principles are:

- Apply data minimisation: reflect whether a full dataset is necessary to achieve the desired outcome or some personal data can be omitted (see 4.4).
- Secure data processing activities: apply security techniques to protect personal location to unlawful disclosure (see 4.6).

5.4. Working with private third parties: exchanging personal location data

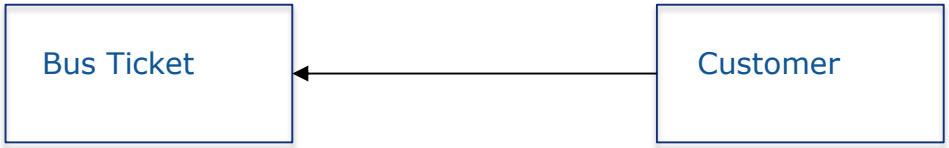
5.4.1. Context

A public administration processes personal location data in the context of its legal obligations. For the execution of this task, it appoints an external private partner to support the processing and storage of personal location data.

5.4.2. Challenges and risks: disclosure of personal location data to third parties

Public administrations increasingly rely on private organisations to deliver their digital services. This includes cloud providers, mobile operators, phone manufacturers and so on. To show this inter-reliance let's look at an example of an everyday occurrence of buying a bus ticket [14]

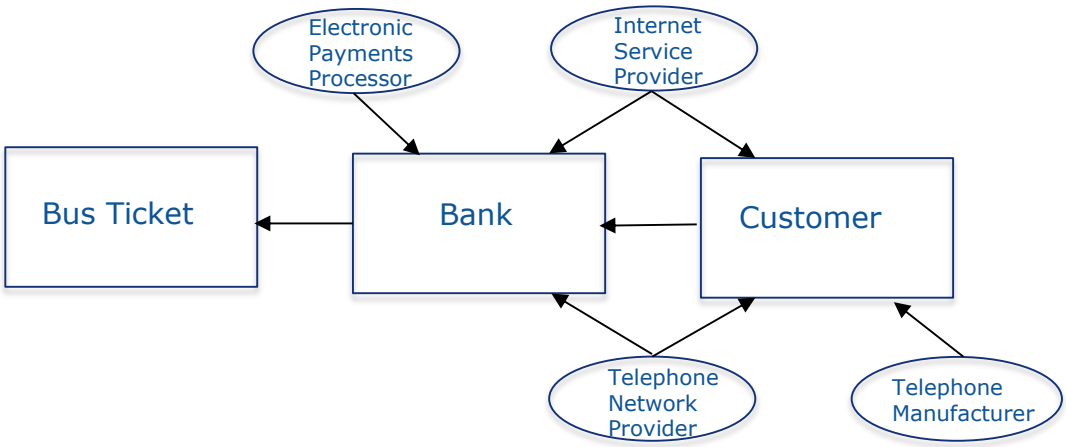
Figure 3: *Cash Ticket Purchase*



In Figure 3 we see that the customer buys the ticket with cash and has a direct one-to-one relationship with the bus company with relatively little or perhaps no personal information being exchanged. The customer is completely anonymous.

In Figure 4 it can be seen how layered, complex and interrelated the process becomes when it goes on line. For online payments, the bus company relies on, for example, internet service providers, the telephone network provider, telephone manufacturers, the telephone operating system. Similarly, the Bank relies on the same network providers as well as a third party gate way system to process electronic payments.

Figure 4: *Credit Ticket Purchase*



A simple single directional transaction with a direct relationship and very limited if any personal detail transferring becomes a complex transaction with the customers details shared by multiple organisation most of which will be in the private sector.

This example shows that the public sector relies on the private sector for the universal delivery of digital services (localised, specialist or in-house services may be different).

The reason for this reliance ranges from reach, expertise, cost, reliability, resources and/or technology availability. As third parties take part in data processing activities, the following points should be kept in mind as they will help reduce any risks, creating both a check list of actions for discussion with third parties and highlight key points for the terms and conditions of the service that need to be included. In addition, it helps form the rationale for a clear and transparent explanation of the service and the response point for queries. (for another example in the transport sector see Annex 1 and the Oyster Card case study).

The actions an organisation can take to mitigate against this exposure include:

- Make sure privacy risks are identified and mitigated using technical and organisational controls. To support this process, a public administration should perform a data protection impact assessment;
- Contractual agreements in which clear responsibilities are defined should be established between the parties involved. These responsibilities describe, amongst other things, non-disclosure, notification processes, right to audit, required level of security and liabilities;
- The necessary legal permissions for the public administration or private third party should be identified and obtained;
- Appoint someone responsible for data protection within every party involved. This person should have sufficient competences and experience and act as a single point of contact amongst the parties involved and the citizens.

5.4.3. Privacy principles

For this specific example the key privacy principles are:

- Perform periodic risk assessments: risks should be identified from two points of view: (1) the privacy risks citizens are exposed to; and (2) the risks the public administration is facing. Note that non-compliance to a legal requirement should be treated as a risk. (see 4.5)
- Appoint a responsible person for data protection (see 4.1)
- Notify data breaches to data subjects and relevant bodies: assigned third parties must notify a public administration in case of a data breach; a public administration should notify the supervisory authority and, in some cases, its citizens. (see 4.8)

6. Recommendations

Section 4 describes the legal obligations related to the collection and processing of personal (location) data. Section 5 elaborates on some specific location data scenarios and examines the related challenges and risks. Insights gained from both chapters are used to define a set of effective and practical recommendations that allow public administrations to go the next level in their quest for adequate personal location data protection. In order to obtain the best possible set of recommendations, these insights are supplemented with the expertise of cyber security and data protection professionals and grounded from discussions with stakeholders on what is most relevant for them. Moreover, all recommendations take into account the specific characteristics identified at the start of this document.

The main high-level recommendation is to start by changing the view and language around who owns the data. Basically, public administrations should move from a mindset of 'owning' personal data to one of being a 'steward' of the data.

The all-encompassing practical recommendation would be to set-up a personal location data protection programme, which goes beyond complying with applicable laws and regulations. This personal location data protection programme can be part of a bigger personal data protection programme. Depending on the amount of (personal) location data an organisation processes, it can pay more attention to location data privacy.

It enables public administrations to become a privacy-aware organisation with respect to personal data protection throughout all its processes. The recommendations described in this chapter all contribute to this idea.

6.1. Set up governance structure for location data protection

Governance is about setting responsibilities and practices with the goal of providing strategic direction, ensuring that objectives are achieved, ascertaining that risks are managed appropriately and verifying that resources are used responsibly. A governance structure provides an answer to the issue of *identity inference* and the fact that location data is *undervalued* (see 1.1).

A location data protection governance structure consists of several building blocks. To achieve an adequate governance structure, the activities below should be completed. Note that these activities are rather generic and can easily be applied to any type of personal data.

- Develop a data protection strategy in line with the organisation's strategy.
- Put together a data protection team and assign responsibilities. One of the most important roles within this privacy team will be the Data Protection Officer (DPO) (see 4.1)
- Implement a policy framework including data protection policies, standards and guidelines. This framework sets out the direction of how to process personal data during its entire data life cycle.
- Define activities with respect to education and awareness, data management (see 6.1), risk assessments (4.5), incident management, and audit and compliance.

- Define metrics to measure to what extent the established data protection programme is effective and adhered to.

Depending on whether location data plays a major role in a public administration's activities, governance activities specific to personal location data could be defined. Some examples of specific governance activities are:

- Guidelines on how to perform self-checks if location data can potentially identify an individual;
- Minimum set of security and privacy controls specific to location data that shall be applied when processing location data (e.g. how to apply anonymisation or pseudonymisation techniques);
- Risk assessment process and supporting risk assessment tools focussing on threats specific to location data.

By putting in place an adequate governance structure, public administrations mitigate the risk of not having full control of the management and protection of personal location data, which is derived from the increased availability and ways of using of location data.

6.2. Set up a location data management programme

Data Management is the development, execution, and supervision of plans, policies, programs and practices that control, protect, deliver and enhance the value of data and information assets [16]. Requirements in this field have changed significantly with continuous technological developments and the growth of IT services such as Internet of Things (IOT), super-connectivity, Big Data, and inter-reliance with the private sector.

Good data management, which must include location data, is crucial as processes and activities become more inter-functional, activities become more data-driven, technology becomes more complex and location data becomes more prevalent. However, not only does the use of data play a major role, but the exchange of data becomes more crucial as well. To manage the abundancy of location data (see 1.1) and to guarantee adequate data quality, it is important to apply good data management.

All this extra data and techniques that are now available with Big Data, machine learning and AI enable insights about individuals that would not have been possible before. The two quotes below capture the changing relationship we have between location data and personal information:

- '…..vast, continuous reams of highly personal data that represents a near real-time snapshot of an individual's movements, activities at specific locations, relationships and affiliations, political beliefs and even mood (sentiment).'
- '…they may be used to not only reveal things about individuals but also to actively structure their life chances and opportunities.' [13]

This level of potential exposure and highly contextual analysis identifies the importance of a location data management programme. The four essential building blocks: of such a programme are (1) data principles and guidelines, (2) roles and responsibilities, (3) processes and (4) supportive tools. If applicable, these buildings

blocks should be aligned with a broader (public sector) location data strategy or location data management.

Based on the above building blocks, location datasets and their use can be identified. Using these insights, data architectures can be modelled. A next step is the introduction of the concepts of reference data and master data management in order to improve the use and management of (location) data. And a final aspect of good data management is the monitoring and maintenance of data quality.

However, polices, protocols and systems are only as good as the people implementing them and this involves two challenges: understanding and culture. In parallel to the building blocks, change management is an important element in a successful data management programme. Create awareness amongst all involved employees on the necessity and potential of data management. An effective approach is one which is top-down, with data champions appointed to lead the change, accompanied by a strong training programme.

6.3. Data subjects are always the data owners

Data ownership is a complex matter. In theory data is no more than a set of characters, which should be contextualised to have meaning or value. Contextual data constitutes information.

According to A. M. Al-Khouri's paper on data ownership [17], personal data comprises, in its strict sense, only personal attributes, which are owned by a data subject. Data subjects sharing their personal data implicitly or explicitly with data controllers, implies delegation of a 'data processing mandate' towards these data controllers. So personal data might have multiple owners as sharing increases. Moreover, if data controllers rely on data processors for the processing of personal data, these data processors are considered data custodians or data stewards.

'True ownership' however lies with the data subject. As data is shared and information is generated, these data should be subject to scrutiny and verification. The only source able to verify the data and confirm the veracity, the 'true owner', is the data subject.

To underpin the above statement, a data subject has been empowered with a series of rights (see 4.7) to be able to scrutinise and verify his data. Public administrations should acknowledge these rights. Moreover, as location data is often embedded in services or applications (see 1.1), public administrations should always be conscious of the potential collection of location data through their services or applications, and protect it accordingly.

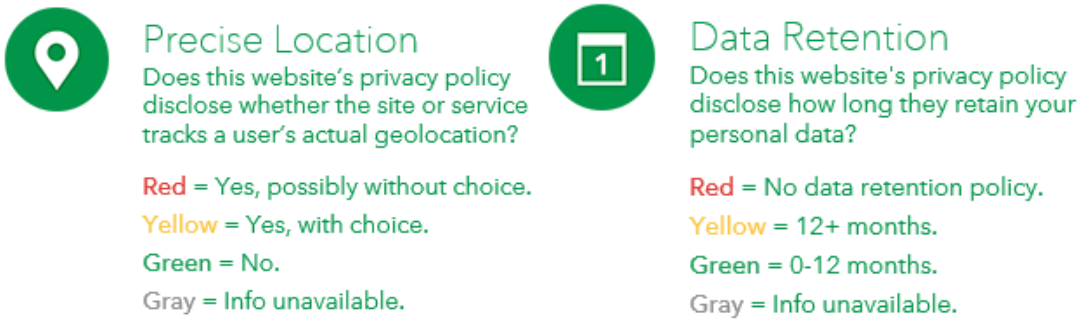
6.4. Create trust through transparency

Data subjects are more inclined to trust you as a data controller or processor if you use an open and transparent communication. As we have seen in the Here survey (Section 5) 68% of those surveyed will share their location information if they know and understand why it is needed and how it will benefit them. Explaining the purpose of your collection, how it will benefit the data subject, the means for safeguarding security or the privacy principles applied in plain and simple text, makes it easier to digest and accept the data processing activity. As location data is sometimes undervalued (see 1.1), it is even more important to make individuals aware of the value of their location data and that this valuable data is protected adequately.

To promote transparency, it is appropriate to establish a privacy contact point, which data subjects can contact for every possible privacy-related question. This contact point can include different communication channels, like a public website, a hotline, an email or even a chat box.

Another means to support transparent and simple communication is the use of standardised icons for data processing activities or controls in place to protect data subjects' privacy. In order for these icons to be effective, they should be comprehensive and evoke similar responses to everyone. Standardised icons should give an overview of the intended processing in an easily visible, intelligible and clearly legible manner (GDPR [2], article 12(7)). Below some examples of privacy icons are depicted to illustrate their use and added value.

Figure 5: *Examples of privacy icons from https://disconnect.me/icons*



6.5. Publish a privacy notice

An important element of a location data protection programme is a privacy notice, sometimes referred to as a privacy policy or privacy statement. A privacy notice describes how an organisation collects, uses, retains and discloses personal data. Unless a (public) organisation's main activity is related to location data, the aspects of location privacy can also be added to a general/organisation privacy notice. Besides a general privacy notice, organisations can also publish application-specific privacy notices and elaborate on how privacy is dealt with within this specific application.

A privacy notice elaborates on the (1) purpose of processing, (2) what personal location data is collected, (3) how the collected data is used, (4) what organisational and technical security measures are in place to protect the personal location data, (5) with whom the personal location data is shared, (6) how a data subject can access or rectify his personal location data, and, not the least, (7) the contact information of the responsible DPO.

As location data is often an *embedded* or simply a *necessary* element in a lot of services or applications (see 1.1), public administrations should explicitly mention the collection and use of (personal) location data in their privacy notice. Moreover, it is important to keep data subjects informed of any changes to the processing of personal (location) data, which should be reflected in the privacy notice.

6.6. Do not confuse privacy and security

Many equate privacy to security, but privacy goes beyond security. Security is just one element for achieving privacy. The Organisation for Economic Co-operation and

Development (OECD) has broken privacy down into eight principles[11], of which security is one component. This security component aims at minimising risks of data loss, unauthorised access, destruction, inappropriate use and improper modification.

The above risks should be weighted and mitigated accordingly. In case of location data, if these risks are insufficiently managed, they might lead to *identity inference* (see 1.1). In order to identify the existing privacy risks, a privacy risk assessment should be carried out (see 4.5).

6.7. It's only as secure as the weakest link

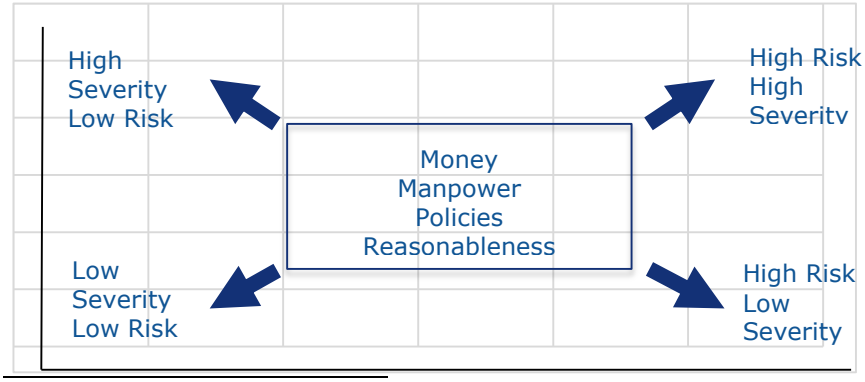
Applying security safeguards is one of the main principles to achieve adequate data protection. When securing data processing activities, security best-practices must be considered. The security principle 'securing the weakest link' states that a solution is only as secure as the weakest link in the whole series of data processing activities, e.g. securing a central information system with rigorous security controls while local end-points are largely ignored, is a waste of money. Every link of the chain must be secured adequately to have an overall secure solution.

Location data might lead to the risk of *identity interference*. As location data may be *undervalued* or the use of it *embedded* in the service, the chance is real that this type of data is insufficiently protected. To know what safeguards to implement at which phase of the information lifecycle, it's important to identify all security and privacy risks throughout the data processing activities. A Data Protection Impact Assessment (DPIA), as described in section 4.5, can support in this process and should be performed for every new initiative or change to an existing one.

6.8. Reduce privacy risks to an acceptable level

Privacy risks should not be reduced to the absolute minimum at any cost. Internationally accepted risk management standards all dictate to mitigate risks to an acceptable level. This acceptable level is called *risk appetite* and defines the impact and likelihood someone can reconcile with. To know what cost is permissible and what cost is excessive, apply the rule of thumb which says that the cost of security controls should never transcend the cost of the impact of a risk. In case of location data, the risk of e.g. *identity inference* (see 1.1) exists and should be mitigated to an acceptable level. This does however not imply that the remaining risks after mitigation should have been reduced to null.

Figure 6: *A model to assess risk*



[11] The other OECD principles relate to collection limitation, data quality, purpose specification, use limitation, openness, individual participation and accountability. Reference is made to all of these principles in the legal obligations described in Section 3.

Figure 6 above shows four possible quadrants of risk ranging from low risk and low severity, to the riskiest space, where both the risk and the severity are high. The centre of figure 6 shows four criteria that can be used in considering the impact of an action. For example, a 100% risk free decision may result is a service that is un-usable or in-accessible due to the cost. If we think of playground where children jump, run and play and want to ensure no accidents ever happen we maybe only allow one child to enter the playground dressed in protective clothing and supported by two to three adults. Of course, this is unreasonable, it would be too expensive to operate, and there would probably not be enough people or protective clothing available – never mind the number of disappointed children left outside. However, it is reasonable to replace the concentrate floor with a sponge base to absorb falls and reduce injuries.

In order to arrive at an acceptable level of privacy risk, Privacy Enhanced Technologies or PETs can be applied. According to Article 29 Working Party, these PETs protect privacy by reducing or eliminating personal data or by preventing the undesired processing of personal data. [18] Moreover, PETS not only help achieving compliance with data protection legislation, but also provide support in the protection of e.g. corporate information.

PETs can be communication anonymisers, which conceal online identifiers such as IP addresses, or encryption which hides and protects information. There are many PETs that enable the protection of privacy, but elaborating on these would go beyond the purpose of this study. More examples of PETs can be found on the website of the International Association of Privacy Professionals (IAPP)[12].

6.9. Prepare for the worst

Public administrations should prepare themselves to respond to major data leaks. Location data is sometimes *undervalued* but just as health or financial data, it might cause serious damage to individuals' private life when leaked, whenever their identity is inferred. Adverse effects for both data subjects and public administrations need to be limited in order to protect data subjects' private life and to retain the trust of stakeholders, despite the data leak.

Crisis management is an organisation's *pre-planned* capability to *respond* to and *recover* from crises, such as major data leaks.

Pre-planned refers to the proactive resilience processes and activities that both prevent and mitigate the impact and the duration of a crisis. Such processes and activities include for example risk management, implementing security measures, audit and assurance, and preparing response and recovery plans (e.g. crisis management plan, crisis communication plan).

Response refers to the recognition and activation of the prepared plans by higher management and taking actions to contain the crisis and mitigate damage.

Recovery involves dealing with the long-term effects of a crisis and how to return to business as usual.

[12] https://iapp.org/news/a/2008-05-introduction-to-privacy-enhancing-technologies

7. Conclusion

Location data can range from typical address data, data over security and traffic control footage to IP addresses or GPS coordinates when citizens access an app or websites. At first glance, this data does not appear to be personal data nor it seems possible to directly identify an individual through this location data. However, although location data might not explicitly reveal an individual's identity, by aggregating disparate data it may be possible to infer an individual's identity.

Personal location data differs from general personal data in some other respects as well. Location data is regularly used by public administrations, but in most cases not as the core of a provided service. Sometimes it is used as necessary functionality, sometimes it is embedded and used unknowingly. And most importantly, location data is sometimes wrongly undervalued when compared to financial or health data. Location data not only says where an individual is, it says who he is and what his interests or preferences are.

As with personal data, the existing general data protection regulations and principles are applicable to personal location data, but there are no specific European legal obligations on location data privacy. As the use of location data presents complications beyond those typically found with respect to other (more obvious) personal data, this document attempts to explain these complications through a series of scenarios and provides specific guidance on the use and protection of personal location data.

By describing some specific scenarios in which location data is used, this guideline creates awareness amongst public administrations on their use of location data. The applied approach for the scenarios can easily be transferred to scenarios specific to a public administration.

The recommendations in this guideline target personal location data. The reasoning behind this focus is because of the characteristics peculiar to location data - specific guidance adds value to the secure and privacy-aware processing of personal location data.

To conclude, it is important to point out the scope limitations applied for this guidance document. The existing legal framework and available best practices for the protection of personal data are elaborate. This document deliberately keeps the processing of personal location data in the context of law enforcement and national security out of scope. However, as mentioned in the introduction, there are numerous applications where location data is used in these contexts. Therefore, further research is considered advantageous and would provide practical guidance to public administrations on how to protect location data in these cases.

References

[1] Toyama, Kentaro, Ron Logan, and Asta Roseway. "Geographic location tags on digital images." *Proceedings of the eleventh ACM international conference on Multimedia*. ACM, 2003.

[2] *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC,* 2016.

[3] American Institute of Certified Public Accountants, Inc. and Canadian Institute of Chartered Accountants., "Generally Accepted Privacy Principles," August 2009. [Online].Available: http://www.aicpa.org/InterestAreas/InformationTechnology/Resources/Privacy/GenerallyAcceptedPrivacyPrinciples/DownloadableDocuments/GAPP_BUS_%200909.pdf. [Accessed January 2016].

[4] *Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE),* 2007.

[5] Organisation for Economic Co-operation and Development, "The OECD Privacy Framework,"11July2013.[Online].Available: http://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf. [Accessed 11 January 2016].

[6] *Regulation (EC) No 45/2001 of the European Parliament and of the Council of 18 December 2000 on the protection of individuals with regard to the processing of personal data by the Community institutions and bodies and on the free movement of such data,* 2001.

[7] *Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data,* 1995.

[8] "Privacy by Design," [Online]. Available: http://www.privacybydesign.ca/. [Accessed 20 May 2016].

[9] European Union Agency for Network and Information Security (ENISA), "Privacy and Data Protection by Design - from policy to engineering," 12 2014. [Online]. Available:https://www.enisa.europa.eu/activities/identity-and-trust/library/deliverables/privacy-and-data-protection-by-design. [Accessed 22 02 2016].

[10] The Location Forum, "Location Data Privacy - Guidelines, Assessment & Recommendations,"1May2013.[Online].Available: https://iapp.org/media/pdf/resource_center/LocationDataPrivacyGuidelines_v2.pdf. [Accessed 11 January 2016].

[11] Kitchin R (2014b) Continuous geosurveillance in the 'Smart City'. DIS Magazine, February. [Online] Available at: http://dismagazine.com/issues/73066/rob-kitchin-spatial-big-data-and-geosurveillance/ (accessed 6th December 2018).

[12] Crawford K (2014) The anxieties of big data. The New Inquiry, 30 May. [Online] Available at:http://thenewinquiry.com/essays/the-anxieties-of-big-data/ (accessed 6th Decembr 2018).

[13] Leszczynski A (2015), Spatial big data and anxieties of control. SAGE Journals, August. [Online] Available at: https://journals.sagepub.com/doi/10.1177/0263775815595814 (accessed 6th December 2018)

[14] Open Data Institute, Personal Data in Transport: exploring a framework for the future. [Online] Available: https://theodi.org/article/personal-data-in-transport-exploring-a-framework-for-the-future-report/ (accessed 6th December 2018)

[15] DG CONNECT, "eGovernment indicators for benchmarking eEurope," 22 02 2001. [Online].Available:https://ec.europa.eu/digital-single-market/en/news/egovernment-indicators-benchmarking-eeurope. [Accessed 23 05 2016].

[16] Dama International, The Dama Guide to the Data Management Body of Knowledge (DAMA-DMBOK), Technics Pubns, 2009.

[17] A. M. Al-Khouri, "Data Ownership: Who Owns 'My Data'?," *International Journal of Management & Information Technology,* vol. 2, no. 1, November 2012.

[18] DG Justice and Consumers, "Glossary," 24 09 2015. [Online]. Available: http://ec.europa.eu/justice/glossary/index_en.htm#glossary-p. [Accessed 2016].

[19] TransportforLondon,"Oystercard,"[Online].Available: https://tfl.gov.uk/corporate/privacy-and-cookies/oyster-card. [Accessed 9 May 2016].

[20] Directive (EU) 2015/413 of the European Parliament and of the Council of 11 March 2015 facilitating cross-border exchange of information on road-safety-related traffic offences, 2015.

[21] Council Decision 2008/615/JHA of 23 June 2008 on the stepping up of cross-border cooperation, particularly in combating terrorism and cross-border crime, 2008.

[22] Interoperability Solutions for European Public Administrations (ISA), "sTESTA," 6 4 2016. [Online]. Available: http://ec.europa.eu/isa/ready-to-use-solutions/stesta_en.htm. [Accessed 2016].

[23] European Public Sector Award (EPSA), "The spanish cadastre, an example of open public administration," 13 June 2012. [Online]. Available: http://www.epsa-projects.eu/index.php?title=The_spanish_cadastre,_an_example_of_open_public_administration#tab=Project_info. [Accessed 14 March 2016].

[24] Elliot, Mackey, O'Hara and Tudor (2016). The Anonymisation Decision-Making Framework.UKAN[Online]Available:http://ukanon.net/wp-content/uploads/2015/05/The-Anonymisation-Decision-making-Framework.pdf [Accessed 6th December 2018]

[25] Information Commissioner's Office (ICO) (November 2012). Anonymisation: managing data protection risk code of practice [Online] Available: https://ico.org.uk/media/for-organisations/documents/1061/anonymisation-code.pdf [Accessed 6th December 2018]

List of abbreviations and definitions

Abbreviations

EULF	European Union Location Framework
ISA	Interoperability Solutions for the European Public Administrations
PNR	Passenger Name Record
GDPR	General Data Protection Regulation
IP	Internet Protocol
GPS	Global Positioning System
DPO	Data Protection Officer
OECD	Organization for Economic Cooperation and Development
ENISA	European Union Agency for Network and Information Security
PbD	Privacy by Design
DPIA	Data Protection Impact Assessment
PIA	Privacy Impact Assessment
ISF	Information Security Forum
NIST	National Institute of Standards and Technology
PET	Privacy Enhanced Technology
IAPP	International Association of Privacy Professionals

Definitions

Anonymisation	The processing of personal data in such a way that the data does no longer relate to an identified or identifiable natural person. [2]
Data controller	The natural or legal person, public authority, agency or any other body which alone or jointly with others determines the purposes and means of the processing of personal data; where the purposes and means of processing are determined by Union law or Member State law, the controller or the specific criteria for his nomination may be designated by Union law or by Member State law. [2]
Data processor	The natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller. [2]
Data subject	A data subject is an identified or identifiable natural person. 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. [2]
Location data	Any data with an implicit or explicit geographic or geospatial reference, ranging from address data to radio signal-based triangulation or IP address location, including data published under the INSPIRE Directive [4]..
Personal data	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. [2]
Personal location data	Any location data directly or indirectly linked to a living individual or that can be directly or indirectly used to identify a living individual.
Privacy	The rights and obligations of individuals and organizations with respect to the collection, use, retention, disclosure and disposal of personal data. [3]
Processing	Any operation or set of operations which is performed upon personal data or sets of personal data, whether or not by automated means, such as collection, recording, organization, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction. [2]
Pseudonymisation	The processing of personal data in such a way that the data can no longer be attributed to a specific data subject without the use of additional data, as long as such additional data is kept separately and subject to technical and organisational measures to ensure non-attribution to an identified or identifiable person. [2]

List of figures

List of tables

Annex I – Case studies

This section contains a list of case studies that have been analysed.

I.1. Case study 1: Oyster

I.1.1. Context

Oyster is an electronic ticketing system of *Transport for London* (TfL), the local government body responsible for the transport system in Greater London. The Oyster card is a contactless smartcard which can hold a credit that can be used to travel through London on bus, tram, Tube and several other public transportation services.

Next to the Oyster card, TfL also offers an Oyster photocard, which bears a photograph of the corresponding traveller and enables reduced rates for specific user groups (e.g. children, students, veterans). *Note*: for this case study, we focus on the regular Oyster card.

Travellers can, optionally in most cases, register their Oyster card through TfL's website and benefit from some extra features. A registered Oyster card allows travellers to pay for purchases in a more simplified way, view journey history of the past eight weeks, protect Oyster cards from loss of theft, or receive updates on planned disruptions on their regular routes.

However, registering an Oyster card is required for e.g. purchasing bus and tram passes valid longer than one month or to use the 'auto top-up' feature, which automatically tops up your Oyster card, based on the provided credit or debit card data, when its balance drops below a certain amount.

When registering an Oyster card, TfL collects a traveller's contact details (name, address, email address and telephone number) and Oyster card unique number. The latter enables TfL to provide a traveller with his journey history. When accessing the online account linked to the Oyster Card, TfL also collects the IP address used by the traveller's computer for the purpose of fraud prevention and detection. [19]

I.1.2. Location privacy challenges

The data collected via the Oyster card enables tracking someone's movements throughout London's grid. For privacy and safety concerns, it is very important to secure this data and protect it from unacceptable use. Moreover, if the data were to be combined with surveillance images, it could be a very powerful tracking tool.

Next to the challenge of protecting this valuable location data, TfL faces another challenge. It shares the personal (location) data with a number of third party providers which operate the majority of the administration and 'back office' services for the Oyster card. Some of these third party providers process data outside the United Kingdom, e.g. in the USA. By relying on these third parties and transporting personal data overseas, TfL must be able to guarantee that all data is processed and managed in a secure way at all times, with respect for the applicable privacy legislations and regulations.

I.1.3. Solution

In order to comply with the requirements of the Data Protection Act 1998 and to control and safeguard the personal data associated with Oyster cards, TfL implemented a range of policies, processes and technical measures.

An example of one of these safeguards is the use of data retention policies for the different types of customer data collected by TfL. These retention polices try to find the right balance between customer privacy and business operations. Some examples of applied data retentions periods are listed below.

- Data about individual journeys are kept for 8 weeks, after which the data is disassociated from the traveller's Oyster card.
- Customer name and contact details are stored for 2 years after an Oyster card was last used.
- Debit or credit card data is stored for 18 months.
- IP addresses collected are stored for 13 months.

I.2. Case study 2: EUCARIS (European CAR and driving licence Information System)

I.2.1. Context

EUCARIS is a peer-to-peer data exchange network facilitating the exchange of mobility related information between European countries. This mobility data includes vehicle registration data or driving licence data and the accompanying personal data. All countries registered to EUCARIS can consult the vehicle and driving licence registers of all participating countries through the *National Contact Points* (NCPs). These NCPs are often the Vehicle and Driving Licence Registration Authorities of the participating countries.

EUCARIS facilitates a number of services including a Cross Border Exchange (CBE) service. The CBE service is used to retrieve vehicle-holder-owner information for vehicles involved in traffic offences in the territory of other participating member states. Furthermore, the CBE can be used to verify if vehicles are registered in multiple countries and thus preventing registration fraud.

I.2.2. Location privacy challenges

EUCARIS is a communication platform that is used to transfer personal (location) data but does not store any personal location data. Even though it does not store the personal (location) data, EUCARIS still has to take appropriate measures to safeguard the privacy of the data subjects involved and fulfil its legal obligations. As data is transferred between different member states, it should take into account not only the directives and regulations regarding data protection at a European level, but national data protection laws of the different member states as well.

Another challenge EUCARIS is facing is proper access management to the EUCARIS platform as it has to rely on the NCPs to grant and facilitate access. It is the NCP's responsibility to decide on what instances get access to the data and the means used to propagate this data to these instances. NCPs are not obliged to give third parties access through the secured EUCARIS web application and thus they can use their own

system to propagate the data. This means EUCARIS has no insights on the level of security of that system.

I.2.3. Solutions

In order to protect the rights and freedoms of data subjects, EUCARIS implemented several safeguards. This case study highlights three main safeguards EUCARIS implemented to protect the data subjects' privacy.

Streamlining a host of legal directives and regulations

All of EUCARIS's processing activities happen on a lawful base: some of them are laid down in EU legislation (e.g. the cross-border enforcement directive [20] or the Prüm Decision [21]), others in the EUCARIS treaty or in bilateral or multilateral agreements. Even though EUCARIS is not bound by national regulation, it did consider the possible differences between countries when drawing up the EUCARIS treaty, leaving enough room for national authorities to decide for themselves how the data will be further processed once on national soil.

Applying data minimisation

To avoid any risk of unlawful processing or insufficient data erasure, EUCARIS does not process or store any personal (location) data. It only provides the platform to facilitate the exchange. The NCP of each member country has its own local EUCARIS server on national soil. Via the central EUCARIS server, bilateral connections are set up to the servers in the other member countries. This way, data does not have to be stored in the EUCARIS system and data is only processed when needed.

Securing data processing activities

The exchange of data between member countries is facilitated by the sTesta [22] network. This provides a highly secure exchange mechanism between NCPs. However, the overall security position also depends on the safeguards applied by the NCPs in their exchange with national entities.

Messages sent throughout the system are plain-text XML-messages. As these messages do not leave the protected environment, it is considered secure. Once the data arrives at the requesting NCP, it can be exchanged with subscribed national entities. This exchange must meet the terms laid down in national and European Law. One example of these terms is the requirement of an audit trail within the Prüm Decision [21].

I.3. Case Study 3: Location data in the Spanish Cadastre [23]

I.3.1. Context

The Spanish Cadastre is a public register containing Spanish real property and real estate. It constitutes a complete digital data model for the whole Spanish territory, except for the Basque Country and Navarra. As well as property identification data (like municipality, address or location of the real estate), the cadastre also holds juridical data (like name and address of the property owner), physical data (like land area) and economic data (like the value of the land).

The cadastral data is primarily collected for taxation purposes, but it also serves as an input for various other public administrations. The Spanish Cadastre also made some

of the cadastral data openly available to all citizens in a response to the numerous requests they received, which initially were to be processed in a non-automated manner.

I.3.2. Location privacy challenges

The Spanish Cadastre is a very open system. Many public administrations can not only access cadastral data, but some are allowed to update the data as well. These are mostly people from municipalities and public administrations. The Spanish Cadastre has only limited control over the actions performed by these public administrations. They have to rely on the agreements made with the other administrations.

Another challenge relates to the introduction of the GDPR [2]. At the time of writing this case study, the Spanish Cadastre was waiting for information from the Central Administration of the State in Spain on how Spanish institutions (like the Spanish cadastre) could prepare themselves for the GDPR [2]. According to the Data Protection Authority of Spain, the Spanish cadastre was already partially compliant with the new GDPR, but certain changes had to be made to be compliant with the new Regulation.

As a last point, the Spanish Cadastre is finding a compromise between its duties as a public administration (making data open to the public) and protecting the privacy of the citizens. In this respect, the Spanish Cadastre is investigating whether it is possible to also open up data relating to the value of a property.

I.3.3. Solutions

In order to protect the rights and freedoms of data subjects, the Spanish Cadastre implemented several safeguards. This case study highlights three main solutions the Spanish Cadastre implemented to protect the data subjects' privacy.

Harmonising open data and protection of privacy

To ensure protection of individuals' privacy, the cadastre provides a secure service guaranteeing data privacy whilst fully supporting transparency within public services. Only data, not subject to data protection law, is visible for all citizens (e.g. surface, location, use, shape, boundaries, cartographic representation, and type of constructions …). The cadastre's policy of open access is compliant with INSPIRE Directive, PSI Directive and Spanish law on citizens' electronic access to public services.

Implementing the rights of the data subject

The Spanish Cadastre has implemented procedures to handle requests relating to the rights of the data subject. For the right to rectification, data subjects can request the Spanish cadastre to update their personal contact details. Notaries, municipalities and registries can access and change the data directly, e.g. a notary can change the ownership of a property when the owner of the property has died. They have the right to process these changes, because the reason for these changes results from a demand by the data subject or a lawful beneficiary (in the event of death). Afterwards, a civil servant of the cadastre will validate the changes, finalising the procedure.

Balancing data protection and open data

All Spanish citizens can access the publicly available data via the website of the Spanish Cadastre, but they can also access the data offline by going to the office of the Spanish Cadastre or to one of the cadastral information points. A citizen can identify himself on the website using his Spanish e-ID card. This allows him to also access his personal data (name, address and national Identifier Number of tile holders and cadastral values), cadastral certifications of various types, and consult and download additional data like a sketch by plant, facade photo, etc.

Entities that collaborate with the Cadastre, like a company, can also access the public services described above. In addition, in case it is described by law, they can access also the data of all real estate within its sphere of competence (non-protected and protected data). For this purpose, it is necessary to be registered off line as a "registered user of cadastre". During this offline registration procedure, the Cadastre checks whether the collaborator has a legal ground and determines what data he will get access to: only data that is necessary to fulfil the purpose can be accessed.

Every access to a property is registered in the logs, allowing property owners to view who accessed their data and for what reasons.

Annex II - Anonymised Data

Introduction

Anonymising data is the process of turning personal data into anonymised data or a form of data where personal information is concealed. The anonymisation of personal data reduces the risk of re-identification of a person and is an example of a 'Privacy by Design' methodology. Data anonymisation has been highlighted in the main body of this guidance. The purpose of this Annex is to provide details of the approaches used.

Data anonymisation can be either at an aggregated level (e.g. men between 20-30 living in Europe) or at an individual level (e.g. test results by patient for a drug trial, where for example a unique reference number is used for each patient so as to conceal their identity). When anonymisation is conducted at the individual patient level it is called pseudonymisation. Basically pseudonymisation is just another format of anonymisation.

The purpose of anonymising data is to ensure protection of personal data when making data available. The objective is for an organisation to convert its personal data into anonymised data and disclose it in the anonymised format, so that this does not amount to the disclosure of personal data. Anonymising data is essentially about minimising the risk of re-identification of person from the anonymised data (see Figure 6 'Model to Access Risk). However, this risk can never be fully eliminated and a decision must be made about the level of risk that would be taken by releasing the anonymised data. This is done by looking at and understanding of the data qualities of your own data as well the wider data pool that is available (see Appendix 1 and 2). [24][25]

Practical considerations

The example below is taken from the UK Information Commissioners Office (ICO) code of practice on anonymisation [25] (Appendix 2 Case Study 3 Page 69). It shows how information from a transport operator's 'go-card' can be anonymised. A 'go card' is an electronic smartcard ticketing system. The user holds the go-card up to the reader to "touch on" before starting a journey, and "touch off" at the end of the journey. The cost of each journey is then deducted from the go-card balance. In this example the go-card information was anonymised by:

a) Creating a 'hashed'(scrambled) version of the unique passenger reference number; and
b) Moving from discreet data points to a range of data. In this case the individual age points were amalgamated under age bands.

Also, note that the location information for the 'Start point' and 'End point' did not need to change in this case. It is assumed that the catchment areas for the start and end points have population levels that are high enough, as well as a high enough throughput of people, so that no person can be identified as being the only person in any bracket.

Figure 7: *Anonymisation of Travel information (adopted from UK ICO)*

Go-card no.	Passenger DoB	Start point	End point	Journey time
WT98765G	01/09/1973	Brooks End	Tree Street	17m 45s
WT45678B	18/09/1933	Brooks End	Tree Street	15m 05s

and this:

Hashed* passenger ref. no.	Age band	Start Point	End Point	Journey time
14793X...	35 - 45	Brooks End	Tree Street	18m
23955P...	75 – 80	Brooks End	Tree Street	15m

* a keyed cryptographic hash function such as SHA356

However, the difficulty arises if there are other data sets available that enable re-identification by a third party. For example, if we refer back to the travel example in figure 7 and change some of the parameters we can see how this could happen. Let us assume that the starting point of Brooks Ends no longer services a large population but only a population of say 500. Now for example if we cross reference with the population statistics and find that there is only one person that is in the age bracket '75-80' we are starting to reveal personal data about this person's frequency, destination and duration of their journeys (see Annex I for more details). This clearly presents a great difficulty for any organisation in dealing with data privacy: How do you assess what you do not know? What other information exists that could enable re-identification? It also re-emphasises the tenet within GDPR, of data minimisation – only the minimum amount of data should only ever be collected, used and published.

In the UK, the High Court ruled that the risk of identification must be greater than remote for anonymised data not to be personal data.

Any organisation disclosing anonymised data should assess whether a person could be identified from the data being released, either in itself or in combination with other data. The standard of 'identification' in this case means establishing a particular connection between the data released and a known individual. An extreme example is if data was released that the income from certain island was €100k and I am the only inhabitant of the island then linking the anonymised income data to population statistics will quickly lead to the identification of an individual.

This amalgamation of different pieces of information is referred to as a 'jigsaw attack' i.e. putting together different pieces of information to create a complete picture of someone. At the simplest level, it can occur if one individual or group already knows another person and can piece the parts together. In the travel example above, if I live in the town or I am a neighbour of the only person who is in the 75-80 bracket then I can use this knowledge to re-identify the person from the anonymised data. As part of the process of releasing anonymised data we need to conduct what is commonly called a 'pen' test to detect and deal with re-identification vulnerabilities. This would involve attempting to re-identify a person from anonymised data and/or other data

sets. An added layer of complexity is that this is an ongoing process, as new or other data could become available that undermines the current anonymised and released data.

In terms of location data, the objective should be to achieve the maximum level of detail that is useful (both in the sense that it does what it is needed for and in the sense that it does not lead to inaccurate or wrongful conclusions) and producing data that is balanced against the protection of an individuals' privacy. As can be seen from Figure 8, making mapping data anonymised can potential significantly deteriorate the level of location data that is available.

Figure 8: *Anonymising Map Data*[13]

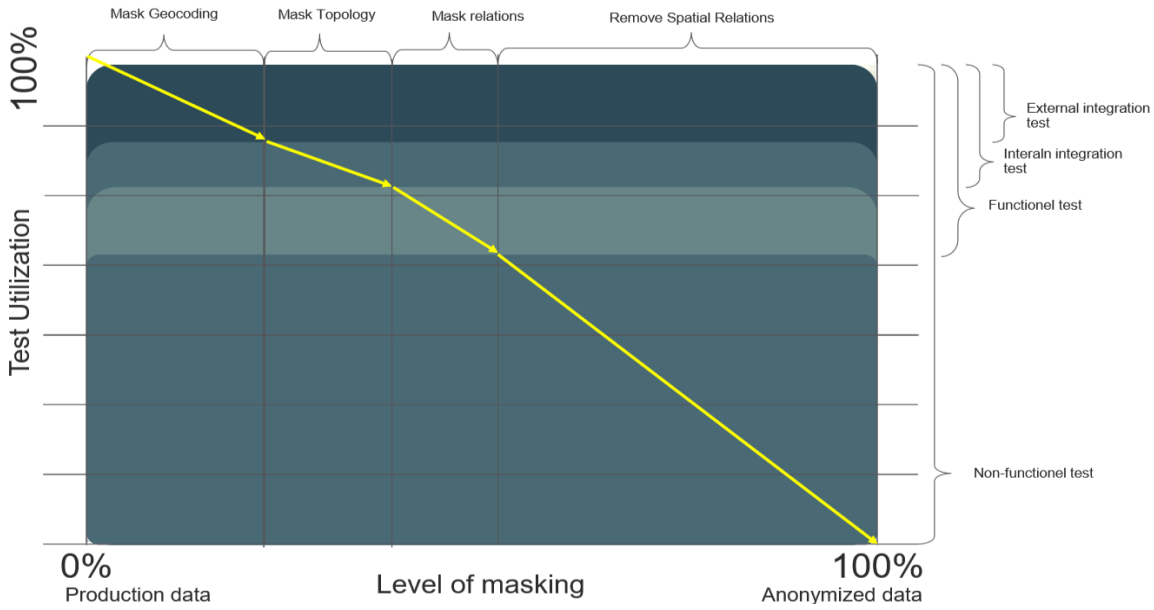


Figure 8 demonstrates a potential understanding of the relationship between anonymisation and mapping data and the resulting level of utilisation and functionality. The scale swings between the extremes of:

1) 100% utilisation and 0% anonymisation (top left hand corner of the diagram); and
2) 100% anonymisation and 0% functionality (bottom .right hand corner of the diagram).

A key assumption of this model is that the functionality lessens as the level of geographic information becomes more generalised.

Under this modelling, it is assumed that a geocode is the most granular level of geographic information. If we begin masking or removing the level of exactness then the level of anonymisation increases and the level of utilisation and functionality also decreases.

In Figure 8 above we may start in the top left-hand corner with geocoding at say room level - with IOT this is the starting point. We then begin the process of

[13] Diagram reproduced with permission from KMD from Denmark

anonymisation and move to geocoding at address point level, then building level, then building in groups of 12 or so buildings, then all the buildings on the street and so on.

What this Figure shows is that as the masking of the geocode information increases the functionality decreases. If we move to mapping topology and spatial relations and do the same exercise it produces the same results - at an even quicker pace until all possible relationships are removed and the data is 100% anonymised and of very limited use for geographical analysis.

This model shows two extremes from 100% useful and functional and 0% anonymised at one extreme to the opposite of the spectrum and 100% anonymised and 0% useful or functional. There are numerous points of data production in between but they carry with them the possibility or risk of re-identification.

In a world of Big Data when unrelated data sets can be combined to gain new insights creating and publishing anonymised geographical data needs some careful consideration and testing before publication.

In general terms, anonymising location data may for example simply mean outputting statistics at a town level rather than building or road level (see Anonymisation techniques below)

There are two broad categories of use of data, namely location data for publication purpose and for internal organisation purposes. Potentially, some of the ways of dealing with the danger of a jigsaw threat when publishing spatial data are as follows:

1) Increase mapping to cover more properties – as mentioned above this may mean outputting data at a town level and not individual house or street level;
2) Reduce the frequency or increase the time period of a publication so it includes more events and is harder to identify a recent or single case;
3) Use formats such as heat maps which provide an overview and do not enable the inference of detailed information about a particular person or small geographically area, for example a small group of houses;
4) Purpose Limitation – Limit the purpose for which the data can be used;
5) Prohibit the use of the data for any attempt at re-identification;
6) Include contractual penalties if recipients breach the terms of use of the data;
7) Adopt measures for the destruction/removal of any accidentally re-identified personal data;
8) Be transparent about why data is anonymised, provide some general information about the way it is done, what are the risks, how the risks were assessed, and what safe guards are in place;
9) Conduct a Privacy Impact Assessment of the process;
10) Have a joined-up approach across the organisation when releasing the anonymised data.

If location data is being used for internal organisational purposes then the following steps can be taken to mitigate the personal data risks:

1) Staff confidentiality agreements;

2) Encryption and key management to restrict access to the data thereby limiting the exposure and risk;

3) Arrangement for the return or destruction of data on the competition of a project. This is particularly important for external consultants who may wish to retain the data for any follow up questions post the termination of the project;

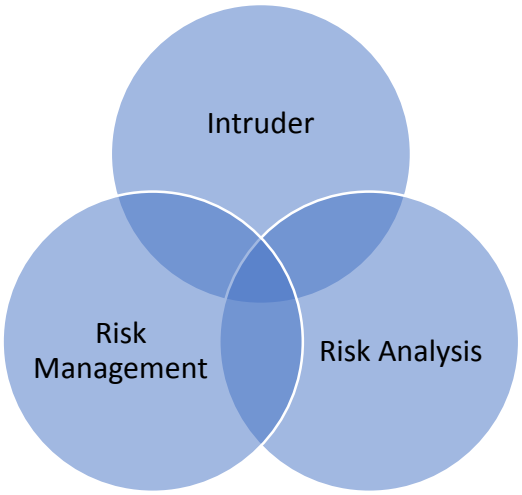
4) Contractual penalties if recipients breach the terms of use of the data;

5) Restrictions on the disclosure of the data.

In summary, anonymisation is essentially a process of risk management which is context-dependent, wherein you must consider the location data and its environment as a total system. Based on this review, a decision can be made as to the location level at which anonymised data could be published. However, the anonymisation of data is a process rather than an end game and an ongoing review of new or amended data format and availability may change the way and level in which anonymised data is published.

Anonymisation decision making framework

Figure 9: *Anonymisation Decision Making Framework (adopted from Anonymisation Decision Making Framework) [24]*



As the diagram above shows there are three distinct aspects:

1) Intruder – There are six distinct elements that need to be considered here;

   a. Motivation of the intruder – what are they trying to achieve?

   b. Means and Resources available to the intruder;

   c. Opportunity – what and how could the intruder access to this and other data sets?

   d. What potential new sensitive information could be learned?

   e. When we think of a potential intruder and how they could use the data have we focused just on one or have we considered it from a variety of different perspectives/angles

f. No data set is perfect, there could be errors or unintended data divergence (e.g. superfluous or inadequate data). It may be useful to consider these issues in terms of critical impact on data privacy and have a programme of checks in place to minimise or remove errors and divergences

2) Risk Analysis – There are two components to consider here;

   a. Attack Type - What computational and statistical methods can be used to attack the data

   b. Data Sources – What other data sources are likely / going to be used? These can/will provide the 'hook' to explore the data.

3) Risk Management

   The results of the analysis are;

   a. Likelihood of an attempt

   b. Likelihood of success

   Based on these results a clear path of what needs to be managed and how to manage the risks should be understood.

The end output of this analysis will point to both the need, level and type of anonymisation (see Anonymisation Techniques, below) that may be required.

In summary, the process is to conduct a data situation audit, review the risks and controls and then look at the potential impact.

Some anonymisation techniques

The techniques below are not shown in order of preference or usefulness or potential risk associated with them. The listing can be used as a guide and not as a definitive set of all techniques available: For a fuller explanation on these techniques and others see Appendix 2 and 3 the ICO document on anonymisation [25]

1) Data reduction – This means the removal of a record or data variable that leads to identifying a person. In location terms it means moving to a more general level of geographical analysis (this may also be called Global Re-coding as it makes the data less specific).

2) Data masking - This involves taking out obvious personal identifiers such as unique building names on a map or a personal name from a piece of information, to create a data set in which no personal identifiers are present. There are a number of variations of this technique, including:

   - Partial data removal – this is a process whereby some personal identifiers are removed such as name and address but others such as dates of birth, remain;

   - Data quarantining – This involves supplying data to a recipient who is not likely to or is unable to have access to the other data that would be needed to facilitate re-identification. It can involve disclosing unique personal identifiers but not the 'key' needed to link these to particular individuals.

3) Pseudonymisation – This is a process of de-identifying data so a reference is attached to a single record that is associated with a single individual without identifying that individual.

4) Deterministic modification – This is similar to pseudonymisation. 'Deterministic' in this context means that the same original value is always replaced by the same modified value. So, if multiple data records are linked, the corresponding records in the modified data set will also be linked in the same way.

5) Micro aggregation – This is a method whereby an observed value is replaced by the average of a group. For example, a heat map rather than a map of individual graded points in an area.

How to obtain EU publications

Our publications are available from EU Bookshop (http://bookshop.europa.eu), where you can place an order with the sales agent of your choice.

The Publications Office has a worldwide network of sales agents.
You can obtain their contact details by sending a fax to (352) 29 29-42758.

LF-NA-28202-EN-N

JRC Mission

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new methods, tools and standards, and sharing its know-how with the Member States, the scientific community and international partners.

Serving society
Stimulating innovation
Supporting legislation

Publications Office

doi:10.2791/420310

ISBN 978-92-79-63495-6