CYBERNETICA

ESTAT 2019.0232 Data Protection Impact Assessment – Scoping Report

Triin Siil, Baldur Kubo, Riivo Talviste, Ville Sokk, Toivo Vajakas, Armin Daniel Kisand

Analysis document

Version 2.9

19.11.2021

137 pages

Y-1440-4

Disclaimer

This document was prepared by Cybernetica AS as part of a procured project under Service Contract No ESTAT 2019.0232 (Ref. Ares(2020)2309804 - 30/04/2020). The opinions expressed in this document are those of the authors. They do not purport to reflect the opinions, views or official positions of the European Commission or its members.

Table of Contents

1. Glossary

1.1. Common abbreviations

Abbreviation	Full definition
DPA	Data Protection Authority
DPIA	Data Protection Impact Assessment
DPO	Data Protection Officer
MNO	Mobile Network Operator
NSI	National Statistical Institute

1.2. Referencing abbreviations

Abbreviation	Full definition
Art	Article
Ibid.	Ibīdem ("in the same place" in Latin)
Op. cit.	Opus citatum or opere citato (“the work cited” in Latin)
P or pp	Page or pages
Rec	Recital
Sec	Section

1.3. Organisations

Abbreviation	Full name
CNIL	French Data Protection Authority (Commission Nationale de l'Informatique et des Libertés)
CoE	Council of Europe
EC	European Commission
ECJ	Court of Justice of the European Union
EDPB	European Data Protection Board
EDPS	European Data Protection Supervisor
ESCB	European System of Central Banks
ESS	European Statistical System
EU	European Union
Eurostat	European Union Statistical Authority, statistical office of the EU, one of the Directorates-General of the Commission
OECD	Organisation for Economic Co-operation and Development
WP29	Article 29 Data Protection Working Party

1.4. Legal acts

Abbreviation	Full title
Charter	Charter of Fundamental Rights of the European Union

General Data Protection Regulation (GDPR)	Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance) – Internet: https://eur-lex.europa.eu/eli/reg/2016/679/oj (04.04.2021).
Regulation on European statistics (RES)	Consolidated text: Regulation (EC) No 223/2009 of the European Parliament and of the Council of 11 March 2009 on European statistics and repealing Regulation (EC, Euratom) No 1101/2008 of the European Parliament and of the Council on the transmission of data subject to statistical confidentiality to the Statistical Office of the European Communities, Council Regulation (EC) No 322/97 on Community Statistics, and Council Decision 89/382/EEC, Euratom establishing a Committee on the Statistical Programmes of the European Communities (Text with relevance for the EEA and for Switzerland), *OJ L 87, 31.3.2009, p. 164–173 (BG, ES, CS, DA, DE, ET, EL, EN, FR, GA, IT, LV, LT, HU, MT, NL, PL, PT, RO, SK, SL, FI, SV), Special edition in Croatian: Chapter 13 Volume 054 P. 186 – 195.* – Internet: http://data.europa.eu/eli/reg/2009/223/2015-06-08 (04.04.2021). - amended by Regulation (EU) 2015/759 of the European Parliament and of the Council of 29 April 2015 amending Regulation (EC) No 223/2009 on European statistics (OJ L 123, 19.5.2015, p. 90–97)
Treaty on the Functioning of the European Union (TFEU)	Consolidated text: Consolidated version of the Treaty on the Functioning of the European Union – Internet: http://data.europa.eu/eli/treaty/tfeu_2016/2020-03-01 (04.04.2021).
Data Protection Directive (DPD)	Consolidated text: 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 - amended by Regulation (EC) No 1882/2003 of the European Parliament and of the Council of 29 September 2003 adapting to Council Decision 1999/468/EC the provisions relating to committees which assist the Commission in the exercise of its implementing powers laid down in instruments subject to the procedure referred to in Article 251 of the EC Treaty.
European Data Protection Regulation (EDPR)	Regulation (EU) 2018/1725 of the European Parliament and of the Council of 23 October 2018 on the protection of natural persons with regard to the processing of

	personal data by the Union institutions, bodies, offices and agencies and on the free movement of such data, and repealing Regulation (EC) No 45/2001 and Decision No 1247/2002/EC (Text with EEA relevance.), PE/31/2018/REV/1, *OJ L 295, 21.11.2018, p. 39–98 (BG, ES, CS, DA, DE, ET, EL, EN, FR, GA, HR, IT, LV, LT, HU, MT, NL, PL, PT, RO, SK, SL, FI, SV)* – Internet: http://data.europa.eu/eli/reg/2018/1725/oj (04.04.2021).
ePrivacy Directive (ePD)	Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector

1.5. Regulatory acts

Abbreviation	Full title
WP29 DPIA Guidelines	Article 29 Data Protection Working Party. Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679. Adopted on 4 April 2017, As last Revised and Adopted on 4 October 2017, 17/EN WP 248 rev.01, p 4. – Internet: https://edpb.europa.eu/our-work-tools/our-documents/guideline/data-protection-impact-assessments-high-risk-processing_en (05.04.2021). Endorsed by the EDPB during its first plenary meeting on 25 May 2018. – The European Data Protection Board. Endorsement 1/2018. – Internet: https://edpb.europa.eu/news/news/2018/endorsement-gdpr-wp29-guidelines-edpb_en (05.04.2021)
EDPS DPIA Guidelines	European Data Protection Supervisor. Accountability on the ground Part II: Data Protection Impact Assessments & Prior Consultation. v1.3 July 2019. – Internet: https://edps.europa.eu/data-protection/our-work/publications/guidelines/accountability-ground-provisional-guidance_en (07.04.2021).
EDPS Preliminary Opinion on Scientific Research	European Data Protection Supervisor. A Preliminary Opinion on data protection and scientific research. 6 January 2020. – Internet: https://edps.europa.eu/data-protection/our-work/publications/opinions/preliminary-opinion-data-protection-and-scientific_en (08.04.2021).
EDPB Document on Health Research	European Data Protection Board. EDPB Document on response to the request from the European Commission for clarifications on the consistent application of the GDPR, focusing on health research. Adopted on 2 February 2021. – Internet: https://edpb.europa.eu/our-work-tools/our-documents/other-guidance/edpb-document-response-request-european-commission_en (07.05.2021).

WP29 Opinion on Purpose Limitation	Article 29 Data Protection Working Party. Opinion 03/2013 on purpose limitation. Adopted on 2 April 2013. 00569/13/EN WP 203. – Internet: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf (14.04.2021).
EDPB Corona App Guidelines	European Data Protection Board. Guidelines 04/2020 on the use of location data and contact tracing tools in the context of the COVID-19 outbreak. Adopted on 21 April 2020. – Internet: https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-042020-use-location-data-and-contact-tracing_en (02.05.2021).
WP29 Opinion on Cloud Computing	Article 29 Data Protection Working Paty. Opinion 05/2012 on Cloud Computing. Adopted July 1st 2012. 01037/12/EN WP196. – Internet: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2012/wp196_en.pdf (10.05.2021).
WP29 Opinion on Anonymisation Techniques	Article 29 Data Protection Working Party. Opinion 05/2014 on Anonymisation Techniques. Adopted on 10 April 2014. 0829/14/EN WP216. – Internet: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf (30.04.2021).
WP29 Opinion on the Concept of Personal Data	Article 29 Data Protection Working Party. Opinion 4/2007 on the concept of personal data. Adopted on 20th June. 01248/07/EN, WP 136. – Internet: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf (01.05.2021).
EDPB Controllership Guidelines	European Data Protection Board. Guidelines 07/2020 on the concepts of controller and processor in the GDPR Version 2.0 Adopted on 07 July 2021. Adopted – After public consultation. – Internet: https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-072020-concepts-controller-and-processor-gdpr_en (24.08.2021).

1.6. ESS common quality framework documents

Abbreviation	Full title
ESCoP	European Statistics Code of Practice. For the National Statistical Authorities and Eurostat (EU statistical authority). Adopted by the European Statistical System Committee, 16th November 2017 – Internet: https://ec.europa.eu/eurostat/web/products-catalogues/-/ks-02-18-142 (12.04.2021).
ESCoP Glossary	Glossary. Defining the main terms used in the European Statistics Code of Practice, as adopted by the ESSC of November 2017, p 1 – Internet:

	https://ec.europa.eu/eurostat/documents/4031688/9439112/Glossary/ (12.04.2021).
QAF ESS	Quality Assurance Framework of the European Statistical System. Version 2.0. – Internet: https://ec.europa.eu/eurostat/documents/64157/4392716/ESS-QAF-V2.0-final.pdf (12.04.2021)
QD ESS	Quality Declaration of the European Statistical System. September 2016. – Internet: https://ec.europa.eu/eurostat/web/products-catalogues/-/KS-02-17-428 (12.04.2021).

2. Introduction

2.1. Abstract

Cybernetica AS developed a proof-of-concept technical solution for a privacy-preserving statistical analysis of mobile location data for Eurostat. The goal was to improve the quality of official statistics by including more accurate data sources and updating the current statistics production methodologies accordingly. Synthetic mobile location data was used to test the solution. Relevant software development, risk assessment and user guide documents were created.

2.2. Background

This document was prepared as a result of a cooperation project between Eurostat and Cybernetica AS[1], which was initiated to:

1) develop a proof-of-concept technical solution in the field of privacy-enhancing technologies for the processing of mobile location data for official statistics ("**Solution**"),
2) test the Solution using synthetically created mobile location data,
3) assess the related privacy risks in the form of a reference instance of privacy impact assessment
(altogether the "**Project**").

Eurostat envisions the Solution to be adopted in practice in the following incremental stages:

1) **proof-of-concept stage** (focus of the Project) – developing, testing and demonstrating the basic functionality of the Solution based on the sample statistical analysis use case in the framework of the reference scenario designed for the Project, using synthetically generated mobile location data.
2) **pilot project stage** – specifying, customizing and running the Solution based on potential statistical analysis use cases identified in follow-up *ad hoc* projects, using real-world mobile location data. This stage falls within the domain of experimental statistics, however, it presumes that Eurostat and relevant national statistical institutes ("**NSI**") have defined concrete needs and requirements for the statistical analysis use cases selected for piloting beforehand. Due to processing real-world mobile location data, a data protection impact assessment ("**DPIA**") and approvals by relevant Data Protection Authorities ("**DPA**") may be required.
3) **production stage** – fine-tuning and implementing the Solution based on previously selected statistical analysis use cases on a permanent basis, using real-world mobile location data. This stage falls within the domain of official statistics. It presumes a clearly defined statistical analysis use case, with a tried-and-tested statistical methodology and privacy risk level assessed in a DPIA and approved by relevant DPAs.

[1] Service Contract No ESTAT 2019.0232 (Ref. Ares(2020)2309804 - 30/04/2020).

2.3. Scope

Eurostat expects the Solution to fulfil certain legal conditions as follows:

- the adoption of the Solution in statistical production process should be compliant with the applicable privacy regulations in Europe, in particular with the General Data Protection Regulation ("**GDPR**"), and hold up to scrutiny by the relevant DPA;
- the Solution will be delivered along with a privacy risk assessment in the form of a reference instance of data protection impact assessment ("**Sample DPIA**"), which would serve as basis and guiding model for a potential partnership between an NSI and a Mobile Network Operator ("**MNO**") in the process of seeking authorisation by the relevant DPA at the national level and/or by the European Data Protection Supervisor ("**EDPS**") for the actual implementation of the Solution for the purpose of producing official statistics.

Eurostat and Cybernetica AS have agreed that the Sample DPIA is delivered in two parts:

a) this present document titled "ESTAT 2019.0232 Data Protection Impact Assessment – Scoping Report" ("**Scoping Report**") serves as the first part of the Sample DPIA;
b) the second part of the Sample DPIA is going to be formalised as a separate document titled "ESTAT 2019.0232 Data Protection Impact Assessment – Evaluation Report" ("**Evaluation Report**") .

The Scoping Report is meant to:

1) fix the scope of the Sample DPIA;
2) analyse the legal issues identified in the course of conducting activities required for carrying out the Sample DPIA activities for the Project ("**Sample DPIA process**");
3) document the decisions made in the Sample DPIA process.

2.4. Intended audience

The document has been drafted for readers with a legal background. In principle, besides the Eurostat staff and contractors, it can be used to help explain the Solution to the Data Protection Officers ("**DPO**") and inhouse lawyers at NSIs and MNOs, as well as officials of DPAs and EDPS.

It is intended to be read and understood as an independent document. However, whenever possible, references to other deliverables of the Project have been added.

2.5. Dependencies

This document should be read and understood in conjunction with the following related deliverables under the Agreement:

1) ESTAT 2019.0232 Solution Analysis
2) ESTAT 2019.0232 Solution Architecture
3) ESTAT 2019.0232 Evaluation Report

3. Goals

MNOs record and store certain types of data for purposes related to delivering telecommunications services (e.g., Call Detail Records collected for billing) and/or in support of network operation (signalling data), which embed information about the (approximate) position and movements of the mobile devices using the MNO network ("**mobile location data**"). Such position and movements of a mobile device can be attributed to a specific customer and/or individual who is using the MNO services ("**Subscriber**"[2]).

Mobile location data is of growing interest to members of the European Statistical System ("**ESS**"). Traditionally, with a few exceptions, NSIs collected data directly from the data subjects via surveys or censuses, where respondents had to provide information about themselves. During the last decade, NSIs have started to extend their scope towards reusing secondary data sources for statistical purposes. First, statistical offices requested access to administrative data (data collected by public authorities) and the EU statistical legislations were amended accordingly.[3] Next, the current trend is to expand the secondary data sources to so-called "big data" that are often collected by private entities – similarly to administrative data sources, it may require explicit reference in statistical legislation to be recognised by potential data holders.[4]

Mobile location data can be used to extract information serving multiple statistical applications and use-cases.[5] Using mobile location data by means of the Solution enables production of different kinds of statistical reports, e.g. about spatial density of present population density and patterns of human mobility. The availability of this kind of technology opens up the possibility to create new types of statistics production processes as well as provide new insights in established statistics production processes. At the same time, new data and technology also bring along new risks, including risks to privacy of individuals.

Previous legal studies have concluded that, in principle, it is not forbidden to use so-called "big data", such as mobile location data, for producing official statistics in the framework of current legislation.[6] At the same time, there is uncertainty regarding legal provisions that explicitly enable the use of such data by NSIs. This uncertainty revolves around the question whether there is an applicable legal basis for NSIs to

---

[2] For ease of reference, we assume the „Subscriber" includes both contractual clients of MNOs and any third parties who use the MNO's network by means of a Subscriber's mobile device.
[3] G. Somers. TASK 3: Legal review Deliverables: D.3.2 Report on legal review covering basic statistical laws and framework legislations D.3.3 Report on legal review covering other legislations. Services concerning ethical, communicational, skills issues and methodological cooperation related to the use of Big Data in European statistics (Contract number 11104.2015.005-2015.799). time.lex, 10 August 2017, p 28. – in the Internet: https://ec.europa.eu/eurostat/cros/system/files/deliverables_3.2_and_3.3_legal_review_final.pdf (04.12.2020).
[4] *Ibid.*
[5] F. Ricciato et al. Towards a methodological framework for estimating present population density from mobile network operator data, p 3. – Pervasive and Mobile Computing. Volume 68, October 2020, in the Internet: https://doi.org/10.1016/j.pmcj.2020.101263 (31.12.2020).
[6] *Op. cit.,* G. Somers, 2017, pp 42-43.

claim mobile location data from MNOs either in the EU law or in the relevant Member State's national law in order to reuse it for the purposes of official statistics.

Even if there is an applicable legal basis for claiming and reusing mobile location data for official statistics purposes, the second area of questions concerns whether this existing legal basis complies with the current requirements of the data protection, electronic communications and statistics law in the EU and the Member States and how these different areas of law interact. For example, what is the status of the legal relationship between the ePrivacy Directive ("**ePD**"), which was adopted during the period of applicability of the Data Protection Directive ("**DPD**") (predecessor of GDPR), and GDPR, which refers to the need to update the ePD.[7] Indeed, a revision of ePD towards a new ePrivacy Regulation ("**ePR**") has been underway for years.[8]

A similar question arises regarding the relationship of the Regulation on European Statistics ("**RES**") and GDPR, as they both contain rules pertaining to processing personal data for statistical purposes.

If no applicable legal basis is to be found, a question arises whether such legal basis should be created and how to make sure it complies with the current requirements of the data protection, electronic communications and statistics law in the EU and the relevant Member States.

For the purposes of the Project, Eurostat expects Cybernetica AS to propose:

1) the most feasible purpose and scope for personal data processing in order to implement the Solution in practice in light of the reference scenario provided by Eurostat;
2) the most feasible approach to support a potential NSI-MNO partnership in pursuing authorization by the relevant DPA to implement the Solution in practice in light of the reference scenario provided by Eurostat.

[7] GDPR Rec 173.
[8] European Commission. Shaping Europe's digital future. Proposal for an ePrivacy Regulation. – Internet: https://digital-strategy.ec.europa.eu/en/policies/eprivacy-regulation (13.05.2021).

4. Executive Summary

The Reference Scenario poses three cumulative requirements to the Solution ("**Core Requirements**"):

1) an MNO applies the change-and-forget method for every single pseudonymisation period (24h),
2) an NSI can produce official statistics from multiple records of the same mobile device over a period of time, which is longer than a single pseudonymisation period (1 year),
3) in order to produce official statistics, the NSI can add confidential calibration data as input to the statistical analysis process.

In order to fulfil the Core Requirements, the following principles were applied when designing the Solution ("**Core Design Principles**"):

a) nobody should be able to see, access or obtain pseudonymous mobile location data through the Solution. This includes, *inter alia*, no extracting of pseudonymisation keys which could be used for reverse pseudonymising the pseudonymous mobile location data at the MNO.
b) the Solution should be able to compute meaningful longitudinal statistical analysis based on the pseudonymous mobile location data,
c) no individual Subscribers should be identifiable from the output results of the Solution.

Essentially, the Core Requirements necessitate that the input data is pseudonymous (pre-processed mobile location data) and the output data is anonymous (official statistics). In legal terms, this means that the input data is personal data and output data is non-personal data. However, there is a strict legal regime under the EU law to protect mobile location data as personal data. In principle, only MNOs are allowed to process mobile location data for providing electronic communications services, as outlined in ePD Art 9. More specifically, ePD Art 9(1) first alternative prohibits MNO from sharing mobile location data with any third parties or otherwise further processing it without the data being "made anonymous". This means that re-use of mobile location data is generally prohibited. Any relevant national laws at Member State level must respect the limitation of ePD Art 9(1) because it is *lex specialis* in relation to GDPR. The exceptions from ePD Art 9(1) first alternative (ePD Art 9(1) second alternative, ePD Art 10(2), ePD Art 15) do not apply in case of the Solution.

According to the Sample DPIA, the most feasible legal route to processing pseudonymous mobile location data by means of the Solution for producing official statistics is to carry out the processing in a single step comprising two concurrent activities ("**2-in-1 approach**"):

1. **making the data anonymous** – the mobile location data is gradually made anonymous.
2. **statistical analysis** – the further processing of mobile location data for producing official statistics is carried out.

The data protection implications of the 2-in-1 approach depend on whether the further processing of pseudonymous mobile location data by means of the Solution qualifies as "made anonymous" under ePD Art 9(1):

a) **if yes**, then the "made anonymous" requirement of ePD Art 9(1) is fulfilled as soon as the pseudonymous mobile location data is encrypted for the Trusted Execution Environment (TEE) within the Solution, considering that, with a sound design and implementation, it is not technically possible for any stakeholder or third party to access any intermediate data other than the final processing results, which are anonymised.
b) **if no**, then the Sample DPIA may need to be adjusted, so as to take better account of the (future) guidelines from the relevant data protection authorities and judgments of the relevant courts.

For the reasons above, the main focus of the Sample DPIA was on the "making the data anonymous" side of the 2-in-1 approach. The Sample DPIA concluded that the further processing of pseudonymous mobile location data by means of the trusted hardware component within the Solution (the enclaves), along with the complementing technical, legal and organisational protection measures applied in the Solution, qualifies as "made anonymous" under ePD Art 9(1), if the Core Design Principles are maintained. This is achieved thanks to a new state-of-the-art introduced by means of the secure computation model used in the Solution, which involves a combination of measures assuring input privacy, output privacy, as well as privacy during processing.[9]

It remains to be seen if the relevant data protection authorities and courts accept the novel interpretations of the concept "made anonymous" as proposed in the Sample DPIA in the context of privacy enhancing technology where no data is shared out of the data owner's organization. If yes, then the following analysis applies:

1) The Solution functions as a condition for the further processing[10]. This means that the condition of "making anonymous" under ePD Art 9(1) holds if and only if all Core Design Principles are maintained. This can be further ensured by applying additional organisational and legal and technical protection measures beyond the Solution.
2) "Making anonymous" in terms of ePD Art 9(1) first alternative is a type of personal data processing. Therefore, it needs to fulfil all the requirements of data protection regulations just as any other type of personal data processing, including compatibility with the purposes for which the data was collected (see the conclusions of the compatibility assessment below), a defined controller and processor(s) (see the conclusions of the controllership assessment below) and a legal basis for processing (see the conclusions of the lawfulness assessment below).
3) As a result of the compatibility assessment carried out for the purposes of the Sample DPIA, it was concluded that making the mobile location data anonymous by means of the Solution for further processing for the purpose of producing official statistics is compatible further use in terms of GDPR Art 6(4) and Art 5(1)(b), because the Solution qualifies as appropriate safeguards in terms of GDPR Art 89(1).
4) As a result of the controllership assessment, it was concluded that either the MNO or the NSI can be designated as the controller in case of making the

[9] For a detailed analysis, please refer to Sections 8.2.4.i) and ii).
[10] See Section 8.2.5. of the Scoping Report.

mobile location data anonymous by means of the Solution for further processing for the purpose of producing official statistics. This depends mainly on whether there is a legal obligation for the MNO to carry out the processing in question. If there is such a legal obligation for the MNO, it presumably determines the MNO as a controller, possibly jointly with the NSI. If there is no such legal obligation for the MNO, the processing must rely on a contractual arrangement between the MNO and NSI and thus presumes a consent from Subscribers as a legal basis for processing. In such case, the NSI and the MNO can agree in the contract that the first acts as the controller and the latter as the processor.

5) As a result of the lawfulness assessment, it was concluded that making the mobile location data anonymous by means of the Solution for further processing for the purpose of producing official statistics can, in principle, be based on consent (GDPR Art 6(1)(a)), legal obligation (GDPR Art 6(1)(c)), public interest/official authority (GDPR Art 6(1)(e)) and legitimate interest of the MNO (GDPR Art 6(1)(f)). A different legal basis may be applied, depending on whether further processing of pseudonymous mobile location data by means of the Solution for producing official statistics is carried out in the proof-of-concept, pilot project or production stage.

6) The next question is whether there are any existing norms in EU or national law which could be relied on as a legal basis for making the mobile location data anonymous by means of the Solution for further processing for the purpose of producing official statistics. Previous legal analysis has shown that such legal basis may exist, for example, in the national law of France[11] and Italy[12]. However, no existing legal basis was identified directly under EU law. Even if there is a potential pre-existing legal basis under EU law or national law of the relevant Member State, it remains to be discussed and analysed whether or not such legal basis can be considered applicable in light of the legal analysis conducted in this Scoping Report. The legitimacy of the relevant EU or national law will be a matter of further legal analysis for each statistical analysis use case selected for implementation in real-world scenarios in the future.

7) The Sample DPIA was carried out in the proof-of-concept stage with the aim of preparing for the pilot project stage. For the purposes of the Sample DPIA, the most feasible approach to processing real-world mobile location data by means of the Solution for producing official statistics in the pilot project *de lege lata* is to:
   (i) accept that it qualifies as “made anonymous” under ePD Art 9(1),
   (ii) select a statistical analysis use case along with appropriate statistical methodologies suitable for implementing in a real-world scenario by means of the Solution,
   (iii) conclude an agreement between the NSI and the MNO for implementing the selected use case, specifying the means and purposes of the processing (the Solution), ensuring protection measures to match the

[11] G. Somers. TASK 3: Legal review Deliverables: D.3.2 Report on legal review covering basic statistical laws and framework legislations D.3.3 Report on legal review covering other legislations. Services concerning ethical, communicational, skills issues and methodological cooperation related to the use of Big Data in European statistics (Contract number 11104.2015.005-2015.799). time.lex, 10 August 2017, pp 29-30, 51-52, 56 – in the Internet: https://ec.europa.eu/eurostat/cros/system/files/deliverables_3.2_and_3.3_legal_review_final.pdf (09.12.2020)

[12] *Ibid.*, p 30 and 57.

requirements for setting up the Solution, dividing the roles of controller and processor, and assistance in obtaining consent from Subscribers,

(iv) carry out a “real” DPIA (taking the Sample DPIA as basis),

(v) consult with or, where necessary, pursue authorisation by the relevant DPA,

(vi) after receipt of the relevant DPIA authorisation, obtain a consent from Subscribers for making their mobile location data anonymous by means of the Solution, where the consent functions as a legal basis for the data processing,

(vii) set up and configure the Solution.

If the relevant data protection authorities and courts do not accept the novel interpretations of the concept “made anonymous” under ePD Art 9(1) as proposed in the Sample DPIA, the only legal route to further processing pseudonymous mobile location data by means of the Solution for producing official statistics is to create a new legal basis for it under EU law or national law of the relevant Member State. Such legal basis may take advantage of the special regime for processing for statistical purposes under GDPR Art 89(2) and other GDPR provisions referring to it, as well as of GDPR Art 11, which allows flexibilities from obligations under GDPR, as long as the controller is able to demonstrate that it is not in a position to identify the Subscribers. This means that even if the 2-in-1 approach applied in the Solution does not qualify as “made anonymous” under ePD Art 9(1), it can still qualify as a set of appropriate safeguards under GDPR Art 89(1). In areas where the ePD does not apply, the Solution can be used as a set of appropriate safeguards for processing other types of data for statistical purposes already today, benefitting from the special regime for processing for statistical purposes under the GDPR.

5. Methodology of the Sample DPIA

5.1. Minimum requirements

A DPIA should contain at least:

1) a systematic description of:
    a. the envisaged processing operations and
    b. the purposes of the processing, including, where applicable, the legitimate interest pursued by the controller;
2) an assessment of:
    a. the necessity and proportionality of the processing operations in relation to the purposes;
    b. the risks to the rights and freedoms of data subjects; and
3) the measures envisaged to:
    a. to address the risks, including safeguards, security measures and mechanisms to ensure the protection of personal data and
    b. to demonstrate compliance with the GDPR taking into account the rights and legitimate interests of data subjects and other persons concerned.[13]

It is up to the controller to choose a more specific methodology for the DPIA, but it should be compliant with common criteria identified in Annex 2 of the Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679[14] ("**WP29 DPIA Guidelines**") adopted by the Article 29 Data Protection Working Party ("**WP29**").

5.2. Approach chosen for the Sample DPIA process

The Sample DPIA was conducted relying on the general outline provided in a DPIA process methodology proposed by a group of data protection researchers and practitioners with extensive experience on the subject ("**DPIA Methodology**").[15] It consists of three stages:

---

[13] GDPR Art 35(7); Article 29 Data Protection Working Party. Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679. Adopted on 4 April 2017, As last Revised and Adopted on 4 October 2017, 17/EN WP 248 rev.01, p 16. – Internet: https://edpb.europa.eu/our-work-tools/our-documents/guideline/data-protection-impact-assessments-high-risk-processing_en (05.04.2021). Endorsed by the EDPB during its first plenary meeting on 25 May 2018. – The European Data Protection Board. Endorsement 1/2018. – Internet: https://edpb.europa.eu/news/news/2018/endorsement-gdpr-wp29-guidelines-edpb_en (05.04.2021).

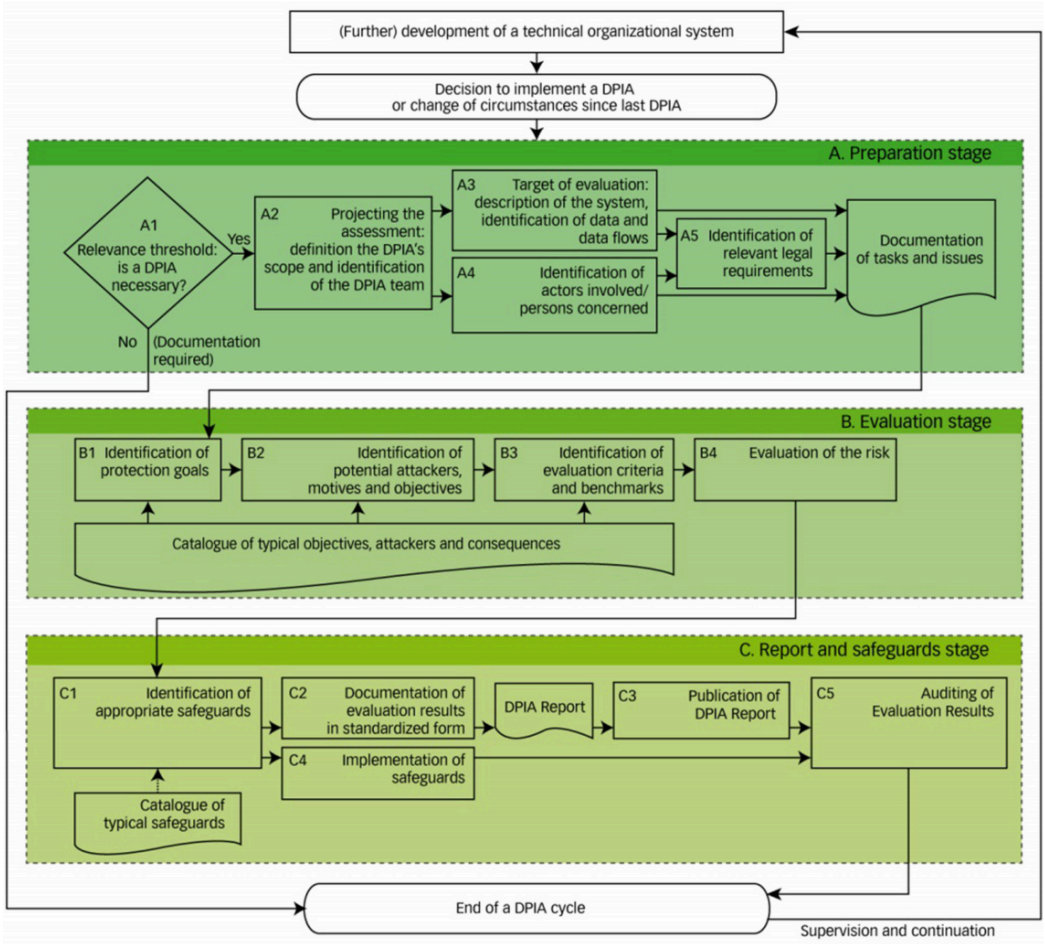
[14] *Op. cit*. Article 29 Data Protection Working Party. Guidelines on Data Protection Impact Assessment (DPIA), 2017, p 17.

[15] The authors suggest a basic DPIA process, which has been derived from the extensive analysis of existing processes and combines procedural as well as evaluation elements, which were tested and approved in practice in the EU projects PIAF and SAPIENT in an extensive empirical assessment of existing PIA schemes. – see F. Bieker et al. A Process for Data Protection Impact Assessment. – S. Schiffner et al (eds). Privacy Technologies and Policy. APF 2016. Lecture Notes in Computer

1) **Preparation stage** – evaluation of a legal obligation to carry out a DPIA, defining the goals and scope of the assessment, choosing the methodology of the assessment, identifying relevant actors involved/persons concerned, identifying relevant legal requirements, documentation of tasks and issues (Scoping Report).
2) **Evaluation stage** – identification of protection goals, potential attackers, their motives and objectives, evaluation criteria and benchmarks, risks.
3) **Report and safeguards stage** – identification and implementation of appropriate safeguards (preparing a plan for risk management), documentation and publication of a report on evaluation results, (Evaluation Report) auditing of evaluation results, supervision and continuation.

According to the DPIA Methodology, a DPIA process is carried out in cycles and the three stages are repeated in each cycle (see Figure 1).

Figure 1 – DPIA process[16]



Science, vol 9857. Springer, Cham., 2016, p 26. – Internet: https://doi.org/10.1007/978-3-319-44760-5_2 (06.04.2021). The paper is also referred to in the European Data Protection Supervisor. Accountability on the ground Part II: Data Protection Impact Assessments & Prior Consultation. v1.3 July 2019, pp 27-28. – Internet: https://edps.europa.eu/data-protection/our-work/publications/guidelines/accountability-ground-provisional-guidance_en (07.04.2021).

[16] *Ibid.*, p 27.

5.3. Approach chosen for the Sample DPIA deliverables

The present Scoping Report is the final deliverable of the Preparation Stage, documenting the issues, decisions made and the justifications relied on. It was drafted between November – April 2021 as a final step before entering the Evaluation Stage and partly in parallel with it. The structure and format of the Scoping Report were modelled after the general guidelines offered in the underlying DPIA Methodology referred to in Section 5.2 above.

A separate document – Evaluation Report – delivers the results of the Evaluation Stage and Reporting Stage. The Report on Evaluation Results was compiled during March-May 2021 based on the CNIL methodology for privacy impact assessments,[17] which is considered a well-known and widely implemented DPIA framework in practice across the EU[18] and was recently updated to meet the new requirements on DPIA introduced by the GDPR.

[17] CNIL. Privacy Impact Assessment (PIA) guidelines. February 2018 editions. – Internet: https://www.cnil.fr/en/PIA-privacy-impact-assessment-en (06.04.2021).

[18] The CNIL methodology for privacy impact assessments is referred to as an example of existing EU DPIA frameworks in the WP29 and EDPS guidelines on DPIA. – See: *Op. cit*. Article 29 Data Protection Working Party. Guidelines on Data Protection Impact Assessment (DPIA), 2017, Annex 1; European Data Protection Supervisor. Accountability on the ground Part II: Data Protection Impact Assessments & Prior Consultation. v1.3 July 2019, p 27. CNIL's security risk assessment methodology, as a critical part of a DPIA, has also been recognized earlier by the European Union Agency for Sybersecurity (ENISA). – see: ENISA. On-line tool for the security of personal data processing. Evaluating the level of risk for a personal data processing operation. – Internet: https://www.enisa.europa.eu/risk-level-tool/risk (27.04.2021).

6. Scope of the Sample DPIA

6.1. Goals

According to the DPIA Methodology, the first step in the preparation stage of a DPIA process is to lay out the goals and scope of the assessment, presuming there is a legal obligation to carry out a DPIA.

DPIA is a process for building and demonstrating compliance with data protection laws.[19] It is envisioned as a tool for facilitating decision-making concerning the processing of personal data and should be started as early as practicable, even if some of the processing operations are still unknown.[20] DPIAs help organisations to ensure data protection by design where it is needed the most.[21]

Carrying out a DPIA is not a one-time exercise, but a continual and iterative process, responding to updates in the development process as well as changes in the risks resulting from the actual processing activity in the implementation phase. For example, if a new technology has come into use or personal data is being used for a different purpose, the revision of a DPIA is not only useful for continuous improvement, but also critical to maintain the level of data protection in a changing environment over time. It is a good practice to continuously review and regularly re-assess a DPIA.[22]

A DPIA can also be useful for assessing the data protection impact of a technology product (e.g., software), where it is likely to be used by different data controllers to carry out different processing operations. In such case, the controller deploying the product remains obliged to carry out its own DPIA with regard to the specific implementation but this can be informed by a DPIA prepared by the product provider, if appropriate. The product provider should share useful information without compromising secrets nor leading to security risks by disclosing vulnerabilities.[23]

In light of the above and considering the current proof-of-concept stage, the Sample DPIA is aimed at assessing the general data protection impact of the Solution when implemented in the context of producing official statistics based on mobile location data. It is a DPIA prepared by Cybernetica AS as the developer of the Solution and provider of the underlying Sharemind HI development platform, keeping in mind that the Solution has potential to be used by different NSIs to carry out different kinds of statistical analysis. The relevant NSIs deploying the Solution in the pilot project and production stages in the future will need to carry out their own respective DPIA with regard to the real-world scenarios and specific statistical analysis use cases at hand. Furthermore, a separate legislative procedure may be needed in order to create a

[19] *Op. cit*. Article 29 Data Protection Working Party. Guidelines on Data Protection Impact Assessment (DPIA), 2017, p 4.
[20] *Ibid.*, p 14.
[21] *Op. cit.*, European Data Protection Supervisor. Accountability on the ground Part II: Data Protection Impact Assessments & Prior Consultation. v1.3 July 2019, p 5.
[22] *Op. cit*. Article 29 Data Protection Working Party. Guidelines on Data Protection Impact Assessment (DPIA), 2017, pp 13-14, 16.
[23] *Ibid.*, p 8.

suitable legal framework to support the deployment of the Solution by an NSI in a real-world situation, unless an appropriate legal basis is already available in the relevant national law. In both cases, this DPIA should speed up the process by offering a basic understanding of the functioning of the Solution and the related risk management considerations with regard to protecting the fundamental rights and freedoms of individuals.

Due to its general nature deriving from the Project being carried out in the proof-of-concept stage, the scope of this DPIA should be defined at a conceptual level rather than in a concrete use case level. The reason for opting for the conceptual level is to reduce the overall complexity deriving from the need to integrate the NSI's statistical analysis process with the MNO's business process and bring it to a minimum acceptable level. The identification of potential statistical analysis use cases suitable for implementing in real-world scenarios by means of the Solution, as well as the design of appropriate statistical methodologies, is a work in progress – the exact use cases and statistical methods remain to be specified as a result of ongoing development efforts by the ESS. Once the choice has been made, the selected use cases along with the accompanying statistical methodologies will need to be subjected to a "real" DPIA, which may be produced taking the Sample DPIA as a starting point.

Nevertheless, there is some initial information available which can be used to sketch out an example of how the Solution may be applied to a close-to-real-world statistical analysis use case in the future. Eurostat has provided this initial information as an input to this Project[24] – it is a simplified version of an actual use case and the accompanying statistical methodology under consideration. Although far from complete, the initial information is helpful in understanding the feasibility and level of risks associated with a statistical analysis use case involving mobile location data. It is a vital starting point for involving specialists form different domains in a discourse concerning the viability and legitimacy of using mobile location data for producing official statistics by means of privacy-enhancing technologies.

For the reasons provided above, there will be some gaps in the documents produced as a result of this DPIA, which will need to be filled gradually as new knowledge becomes available. It is a first step in a way towards demonstrating the technical feasibility and legal compliance of implementing privacy-enhancing technologies in facilitating secondary use of big data from sources outside the ESS, while maintaining the level of data protection and statistical confidentiality required under applicable laws and regulations.

6.2. Context

Before describing the target of evaluation of the Sample DPIA in greater detail, some framework conditions need to be laid out in order to explain the function of the statistical analysis use case and clarify the extent of changes it may bring along compared to the status quo. For example, the purposes of the Project, it is important for the reader to understand the current approach of including a territorial dimension in official statistics in the EU, so as to be able to compare it with the new approach

[24] See: Eurostat. Specification of test use-cases for project ESTAT 2019.0232.

introduced in the Project. This way, the shortcomings of the current approach become evident and provide a practical reason for considering a new approach.

The European Commission has recently made available data for several different territorial typologies across the EU, which has stimulated policymakers to carry out new kinds of policy analyses using a territorial dimension. “Grouping different types of regions and/or areas according to territorial types can help in understanding common patterns, for example, urban areas/regions generally perform better in economic terms and may act as hubs for innovation and education; at the same time, they may also be characterised by a range of different challenges such as congestion, pollution or housing problems.”[25]

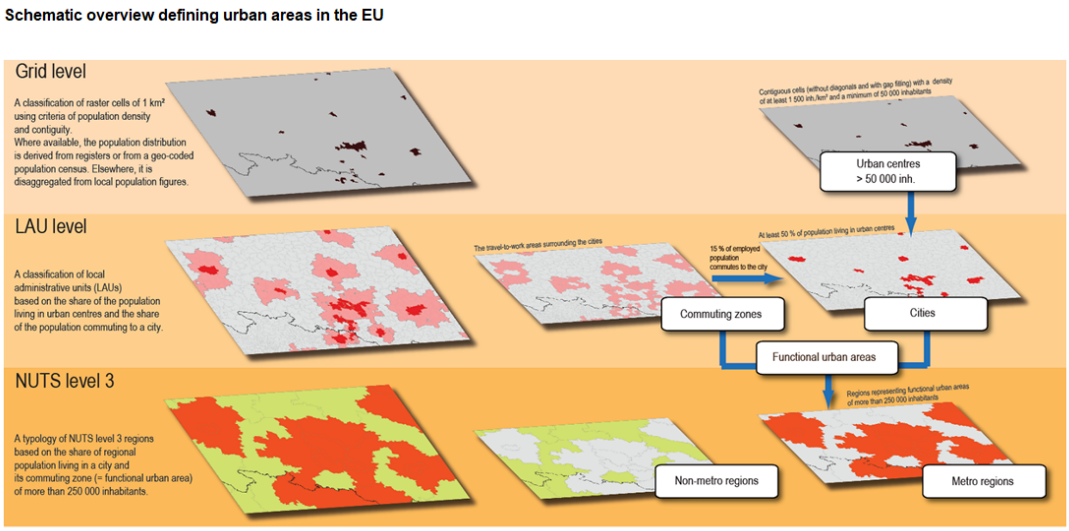
The main territorial typologies can be divided into three different groups:

1) **grid typologies** – Eurostat collects population statistics based on 1 km² grid cells. These are very detailed statistics, which are used to establish various cluster types — namely, urban centres, urban clusters and rural grid cells.[26]
2) **local typologies** – based on statistics for local administrative units (LAU), such as municipalities or communes across the EU. These statistics may be used to establish local typologies including the degree of urbanisation (cities; towns and suburbs; rural areas); functional urban areas (cities and their surrounding commuting zones); coastal areas (coastal and non-coastal areas).[27]
3) **regional typologies** – statistics that are grouped according to the classification of territorial units for statistics (NUTS). They provide information at a relatively aggregated level of detail.[28]

The three different types of territorial typologies are all based on the same basic building blocks – classifying population grid cells to different cluster types and then aggregating this information either by LAU or by region to produce statistics for a wide variety of different typologies.[29] Figure 2 below presents an example for how urban areas in the EU are defined at three different — but coherent — levels:

[25] Eurostat. The Methodological manual on territorial typologies. 2018 edition, p 26. – Internet: https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gq-18-008 (23.04.2021).
[26] *Ibid.*, p 7.
[27] *Ibid*.
[28] *Ibid*.
[29] *Ibid*.

Figure 2 – Schematic overview defining urban areas in the EU[30]



Note: for more information, http://ec.europa.eu/regional_policy/sources/docgener/focus/2012_01_city.pdf
Source: European Commission, Directorate-General Regional and Urban Policy, based on data from Eurostat, JRC, national statistical authorities

eurostat

6.2.1. Population grid

A population grid if one of the three basic building blocks that underpin the various territorial typologies. It is composed of a set of equally-sized cells containing population counts for each cell. Eurostat prefers the use of a 1km^2 square grid that is overlaid across the EU territory.[31]

In practice, the population distribution data underlying the grid level is currently derived from registers or from a geo-coded population census or disaggregated from local population figures. There are three methodological solutions foreseen for establishing the total number of inhabitants living in each of the 1km^2 grid cells:

- **aggregation method** – the preferred method for producing population grid data, which is based on aggregating geocoded micro data. For example, aggregating a geocoded point-based data source, such as an address.[32]
- **disaggregation method** – in the absence of geocoded micro data, there are alternative approaches to producing data for the grid. The first alternative is the disaggregation method, which uses population statistics for LAUs in combination with auxiliary spatial data. The total population count for a LAU may be disaggregated using data on land use and/or land cover to estimate the number of inhabitants that are living in each 1 km² grid cell (e.g. through the visual inspection of satellite images overlaid on the grid to determine if there are any buildings in each grid cell).[33]
- **hybrid method** – based on combining the aggregation and disaggregation techniques, this method provides a compromise between accuracy and the availability of data. “Hybrid solutions may refer to using different source data

[30] *Ibid.*, pp 8-9.
[31] *Ibid.*, p 13.
[32] *Ibid.*
[33] *Ibid.*, p 15.

to establish a geocoded framework, for example, combining geospatial, administrative and statistical sources."[34]

"Population grids are a powerful tool for describing the spatial distribution of a population and are particularly useful for analysing socioeconomic phenomena that are independent of administrative boundaries."[35] Because population grid statistics are detailed in nature, they are considered more advantageous compared to traditional statistics that are based on larger administrative or statistical areas.[36] At the same time, "one negative effect of developing grid-based statistics that have a much greater level of geographical detail is that there are increased concerns around data confidentiality and/or the risk of disclosure. Moreover, when introducing supplementary variables linked to the population (such as analyses by sex, by age or by type of housing) these issues may become even greater."[37]

In 2010, Eurostat and European Forum for Geography and Statistics launched a long-term programme designed to set up and promote the use of geospatial statistics including grid-based statistics through developing a methodology for official geospatial statistics in the EU, both for individual EU Member States and the EU as a whole, including developing a set of common guidelines for the collection and production of population grid statistics.[38] As part of the programme, a standardised population grid – GEOSTAT 2011 – was developed using the aggregation method. GEOSTAT 2011 is based on a 1 km² grid, which has been considered a good compromise between analytical capacity and data protection for European data.[39] Also, the GEOSTAT 2011 population grid only contains information for the total number of inhabitants at their place of usual residence – this statistic was usually considered as non-sensitive by national statistical authorities which, as a result, did not apply any data protection methods for confidentiality issues. However, national laws may require NSIs to protect the identification of individual citizens – where the confidentiality thresholds for GEOSTAT 2011 were established under national laws, the minimum number of inhabitants per grid cells was 3 to 10 individuals; under this threshold the population count was suppressed.[40]

6.2.2. Functional Urban Area

Functional urban area ("**FUA**") is one of the local territorial typologies described above. It consists of a "city" (densely inhabited) and its "commuting zone" (less densely populated) whose labour market is highly integrated with the city.[41] "City" is a LAU where at least 50 % of the population lives in one or more urban centres.[42]

---

[34] *Ibid.*, p 16.
[35] *Ibid*.
[36] *Ibid*.
[37] *Ibid*.
[38] *Ibid*.
[39] *Ibid.*, p 13.
[40] *Ibid.*, p 16.
[41] Eurostat. Statistics explained. Glossary. Functional urban area. – Internet: https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Functional_urban_area (23.04.2021).
[42] Eurostat. Statistics explained. Glossary. City. – https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:City (23.04.2021).

"Commuting zone" contains the surrounding travel-to-work areas of a city where at least 15 % of employed residents are working in a city.[43]

FUAs do not cover the whole territory of a country but rather the more densely populated areas. Defining FUAs requires three steps ("**FUA methodology**"):

1) **classifying urban centres** by means of densely populated grid cells;
2) **classifying cities** by covering the urban centres with LAUs;
3) **classifying commuting zones** based on commuting patterns:
    a. if at least 15 % of employed persons living in one city work in another city, these cities are treated as a single destination for the commuting analysis;
    b. all LAUs from which at least 15 % of the employed population commute to the city are identified as commuting zones;
    c. LAUs surrounded by a single functional urban area are included as part of the commuting zone and non-contiguous LAUs are excluded from commuting zones.[44]

FUA as a type of classification is linked to other classification types, such as the degree of urbanisation and typology for metropolitan regions. It is used as a basis for the city statistics data collection.[45] Currently, there is no EU legislation on the collection of city statistics and they are provided on a voluntary basis only.[46]

6.3. Target of evaluation

According to the DPIA Methodology, the next step in the preparation stage of a DPIA after providing the goals of the assessment is to describe the target of evaluation, which defines the scope of the DPIA. "It is paramount that the controller is aware of the extent of the processing operations in order to determine how these may affect the rights of the individual. […] A DPIA must describe the predefined object of evaluation in its entirety, including in technical as well as the organizational implementation at the controller level. This concerns any use cases that are to be implemented and should pay particular regard to the purposes of the data processing."[47]

Based on the current approach to including a territorial dimension in official statistics in the EU (see Section 6.2 above), it is possible to envision alternative population grid statistics and territorial typologies when using mobile location data as a source for evaluating population distribution. Inspired by the concept of FUA (see Section 6.2.2 above), the Eurostat staff has developed, solely for the purpose of this specific Project, a toy methodology based on mobile location data ("**Toy Methodology**"). The

[43] Eurostat. Statistics explained. Glossary. Commuting zone. – Internet: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Commuting_zone (23.04.2021).

[44] Eurostat. The Methodological manual on territorial typologies. 2018 edition, pp 49-51. – Internet: https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gq-18-008 (23.04.2021)
[45] *Ibid.*, pp 59-60.
[46] Eurostat. Cities. Background. – Internet: https://ec.europa.eu/eurostat/web/cities/background (26.04.2021).
[47] *Op. cit.*, F. Bieker et al. A Process for Data Protection Impact Assessment. 2016, p 28.

FUA-inspired concept used in the Toy Methodology indicates an approximation of cities and their commuting zones within the scope of what is realistically achievable by mobile location data – it is hereinafter referred to as Functional Urban Fingerprint ("**FUF**").

The target of evaluation of the Sample DPIA is the further processing of pseudonymous mobile location data by means of the Solution for the purposes of official statistics. The evaluation is conducted at a conceptual level, applying the Toy Methodology (see Section 6.3.2 below) for the reference scenario (see Section 6.3.1) in a proof-of-concept setting (see Section 6.3.3 below) (altogether "**Sample Use Case**") as an example for, not as an object of, the Sample DPIA. The focus of the current analysis is on evaluating the change in the level of risks to the fundamental rights and freedoms of individuals, which may occur due to introducing mobile location data as a new source for computing the territorial dimension in official statistics by means of the Solution.

The Sample Use Case has been designed to test how the current statistics production processes using the FUA concept can be emulated by means of the Solution, if the underlying population grid were derived from mobile location data. Its purpose is to provide new insights to Eurostat for developing the concept of FUF. Therefore, it should be treated as a tool to measure how efficiently the privacy-preserving functionalities of the Solution work and what could be improved. The Sample DPIA is not meant to assure the compliance of the Sample Use Case with applicable laws nor to evaluate the specific security and privacy risks related to the Sample Use Case – this shall be an object of future assessment.

6.3.1. Reference scenario

Eurostat initially provided the following reference scenario for the purposes of the Project:

a) MNOs record and store mobile location data, which can be associated with Subscribers.

b) In addition to the primary purposes for which mobile location data is initially collected, it can also be useful for secondary purposes. For example, MNOs can provide certain value-added services to Subscribers based on analysis of mobile location data. In order to prevent long-term tracking of the Subscribers during secondary use of mobile location data, MNOs apply pseudonymisation to mobile device identifiers and such pseudonymisation is based on a secret key. The key is changed periodically and therefore also the generated short-term pseudonyms are different in each pseudonymization period. As soon as a new key for the next pseudonymisation period is generated, the previous pseudonymisation key is deleted – this change-and-forget method helps to decrease the risk and potential impact of personal re-identification for mobile location data over multiple pseudonymisation periods.

c) NSIs within the ESS are considering mobile location data as a potential new data source for the production of future official statistics regarding mobility and presence patterns of Subscribers. In order to reuse the mobile location data

for such purposes, an NSI would need to be able to draw insights from multiple records of the same mobile device over a period of time, which is longer than a single pseudonymisation period.

d) The Solution should enable the extraction of statistics based on long-term analysis of mobile location data that are pseudonymised with short-term pseudonyms, while not increasing the level of risk and potential impact of re-identification of Subscribers beyond the time interval covered by a given short-term pseudonym.

e) The Solution should also enable the possibility to add non-public input data from the NSI to the computation of statistics for calibration purposes.

f) Compared to the change-and-forget method, the autonomous capability of the MNO to process and extract information from the pseudonymised data (e.g. for providing value-added services) should neither be increased nor decreased by the adoption of the Solution. The main goal is for the Solution to have a clear potential for practical adoption in real-world scenarios, particularly for what concerns technical feasibility and legal viability.

6.3.2. Toy Methodology

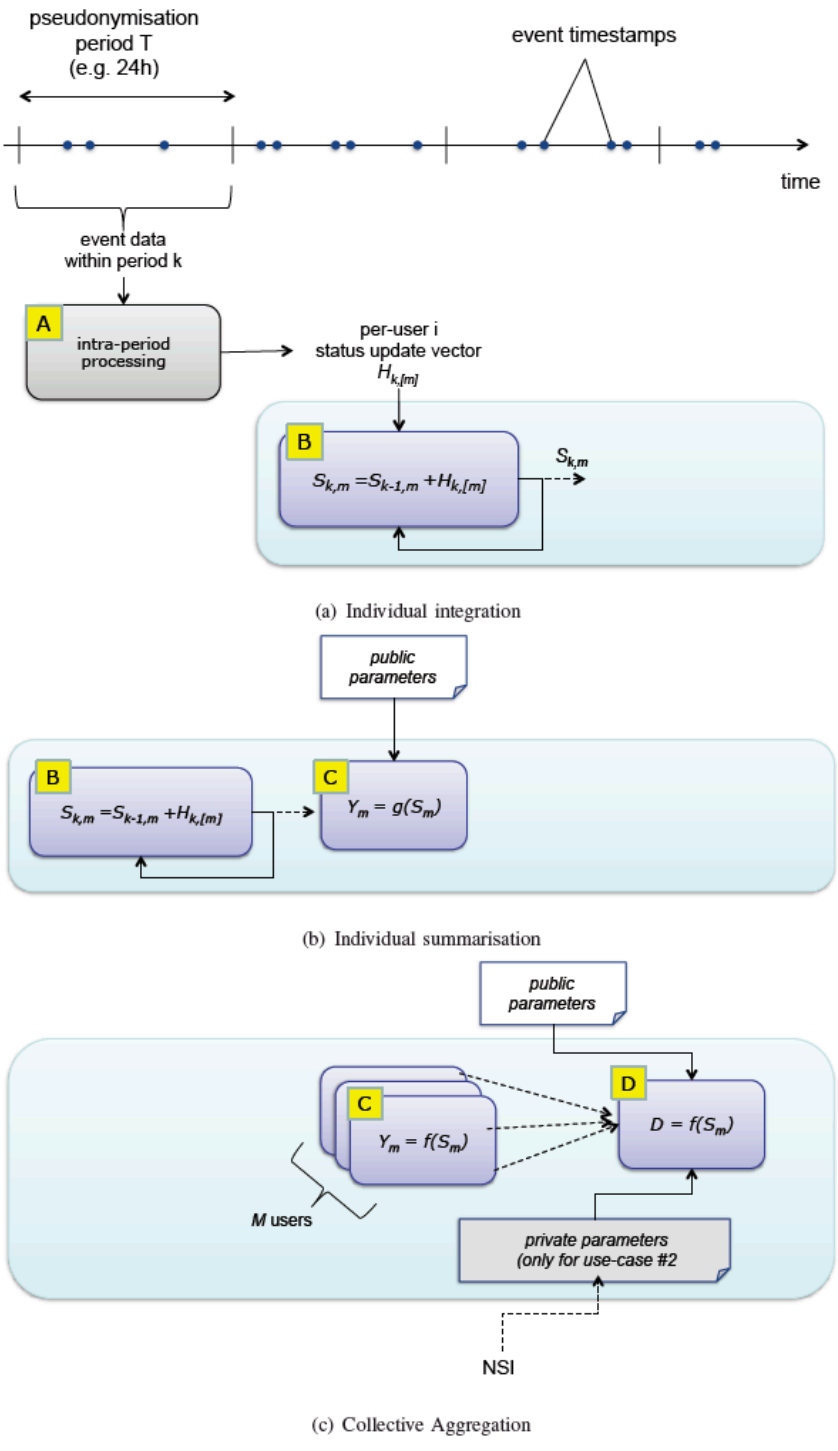
In addition to the reference scenario, Eurostat later provided the Toy Methodology for the purposes of the Project. The Toy Methodology is divided into a number of technical stages, which are described in detail in the technical specification provided by Eurostat.[48] These stages follow a step-by-step approach, whereby the pseudonymous mobile location data are gradually de-identified, each next step reducing the link between the mobile location data and the relevant Subscriber until its elimination. All these activities can be observed in two levels:

1) "**Meta Level**" – the structural elements of the Sample Use Case (Module A, Module B, Module C and Module D), which will most probably be preserved in the official methodology used in the statistical analysis use cases implemented for real-world scenarios in the future;
2) "**Use Case Level**" – the dynamic elements of the Sample Use Case, which can differ from one statistical analysis use case to another, when implemented for real-world scenarios in the future (number of countries, amount of Subscribers, content of public and private parameters, specific computation algorithms, choice of SDC techniques etc).

The technical stages can be relied on also for the purposes of the legal analysis in this Scoping Report. In order to facilitate understanding of the Toy Methodology by readers with non-technical background, the technical stages are summarised within the framework of the reference scenario based on the schematic flow depicted in Figure 3 below.

[48] See: Specification of test use-cases for project ESTAT 2019.0232.

Figure 3 – Schematic flow: the private computation blocks are embedded in the cyan box.[49]



i. Individual integration

This stage is divided into two parts:

[49] *Ibid.*, p 9.

a) Pre-processing step (Module A)

The data processing is carried out at the level of an individual Subscriber. Each Subscriber's mobile device has been previously assigned a pseudonym, which changes periodically after each 24h interval. The MNO provides mobile location data concerning cells in a 1kmx1km population grid that pseudonymous mobile devices have visited during each 24h interval.

Module A integrates all visits per each pseudonymous mobile device during each 24h period into a data structure ("**Intra-Period Footprint**[50]"). The Intra-Period Footprint provides a score for each grid cell depending on how often and how long a pseudonymous mobile device spent time in it ("**Footprint Score**"). The Footprint Score is computed from a sequence of visits using a predefined algorithm provided by the NSI (public information).

The pre-processing stage carried out in the Module A is not part of the Solution.

b) Accumulation of individual footprint (Module B)

The Intra-Period Footprints and pseudonyms of mobile devices are provided as input to the Solution.

The first part of the Solution – Module B – links the different pseudonyms associated with the same mobile device across all 24h intervals and adds the Intra-Period Footprints of all associated pseudonyms together. As a result, all visits per each (non-pseudonymous) mobile device over a long period can be summarised in a data structure ("**Longitudinal Footprint**").

ii. Individual summarisation

The second part of the Solution – Module C – takes as input the Longitudinal Footprint for each single (non-pseudonymous) mobile device and selects the grid cells that have been visited more frequently and more regularly by a given mobile device. The output of Module C is a data structure consolidating the Longitudinal Footprints of all single mobile devices ("**Consolidated Footprint**") and, thus, still refers to such individual devices. The Consolidated Footprint has only a handful of non-zero entries (grid cells) and aims at representing the "usual environments" of a generic individual device.

The thresholds for determining most visited grid cells per mobile device will be provided as input to Module C by Eurostat (public information). Based on that, the top ranked grid cells ("**Top Tiles**") will be calculated per each mobile device in Module C, and will therefore represent the estimated usual environment zones for each mobile device.

iii. Collective aggregation

[50] Note the difference between the term "footprint", referring to data structures associated with individual mobile devices, and "fingerprint", for aggregate data structures associated with grid cells or Reference Area.

The last part of the Solution – Module D – takes as input the Consolidated Footprint and computes aggregated statistics over the whole population of mobile devices.

An approximation of cities is given based on a pre-defined list of territories, which roughly correspond to administrative urban areas ("**Reference Areas**"). Each Reference Area is a list of contiguous grid cells. The Reference Areas are provided by the NSI as an input to the Solution (public information).

An approximation of commuting zones is calculated within the scope of what is realistically achievable by mobile location data. In order to estimate the load of movement between a city and its commuting areas, a value is calculated which indicates how strongly a grid cell outside the Reference Area is connected to the Reference Area. This value is related to the share of devices having a Consolidated Footprint intersecting both the grid cell and the Reference Area.

The NSI has the option to provide a secret input (can not be visible neither to the MNO nor to any other third party), which gives the resident count of each grid cell ("**Resident Count**"). The Resident Count has been estimated using a population census and can be used for calibrating the results of the analysis.

Statistical Disclosure Control ("**SDC**") is applied to the results of the analysis as a last step before releasing the output. SDC omits grid cell values that do not exceed a pre-defined threshold, in order to comply with the statistical principles laid out in the Treaty on the Functioning of the European Union[51] ("**TFEU**") Art 338 and Article 2 of the ESR[52]. This threshold is fixed beforehand and cannot be changed after the Solution has been deployed. In order to change the SDC threshold, it has to be changed in the source code, the new source code has to be compiled, the new version has to be deployed and enforcers have to approve it again. The state of the previous version will not be accessible in the newly deployed enclave. The threshold values are made public, along with a detailed description of the SDC method and the associated code.

In the last step, the following reports are made available as a result of the analysis:

1) **Fingerprint Report** – indicates for each grid cell how many mobile devices are typically found in this grid cell at a given time of day.
2) **Population Density Report** – indicates for each grid cell how many mobile devices had this grid cell as their No 1 Top Tile, corresponding to the most likely main place of living.
3) **FUF Report** – indicates an approximation of cities and their commuting zones within the scope of what is realistically achievable by mobile location data.

In addition to the reports, the following aggregate results are reported for quality control purposes:

1) **Highly Nomadic Users** – the number of Subscribers who did not have Top Tiles or where their computation did not succeed.

[51] Treaty on the Functioning of the European Union. Consolidated text: Consolidated version of the Treaty on the Functioning of the European Union – Internet: http://data.europa.eu/eli/treaty/tfeu_2016/2020-03-01 (04.04.2021).
[52] The thresholds are provided and calculated by the NSI in accordance with relevant SDC methodologies.

2) **Observed Total Users** – the total number of observed individual Subscribers.
3) **Adjusted Total Users** – the total number of observed individual Subscribers after the optional calibration.

6.3.3. Proof-of-concept setting

In consultation with Eurostat, the following requirements were established for the proof-of-concept setting:

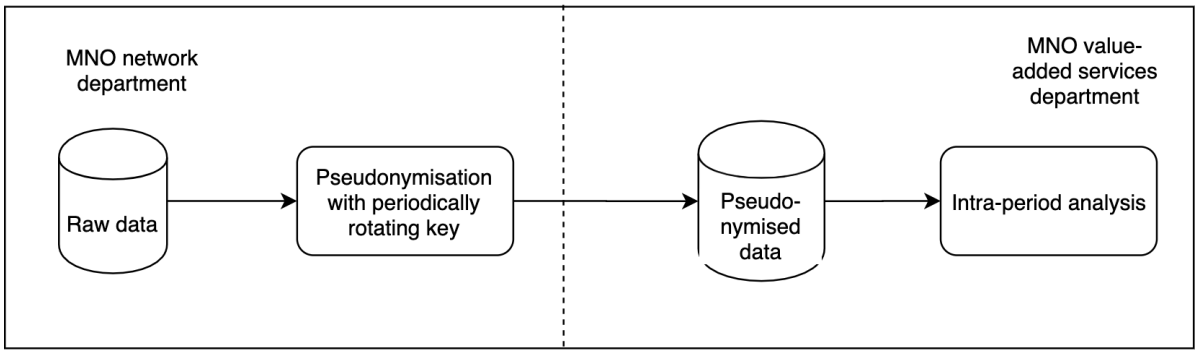
1. The mobile location data used for statistical analysis by means of the Solution is provided by one hypothetical MNO. This is the only source of mobile location data, i.e. data from other sources (e.g. customer relationship management systems) is not used.
2. The mobile network of the MNO covers the full national territory of a hypothetical country, which is an EU Member State, and has approximately 100 000 000 Subscribers.
3. The pseudonymised mobile location data is stored for the purposes of statistical analysis by means of the Solution for up to one year.
4. The statistical analysis by means of the Solution is ordered by one hypothetical NSI, located in the same country as the mobile network of the MNO.
5. The domain of official statistics and the broader production process embedding the Sample Use Case is unspecified.

Even though conducted in proof-of-concept setting, it is presumed that the processing of mobile location data is based on actual real-world personal data. Otherwise, in case of using synthetically generated mobile location data, the relevant data protection laws will not apply.

6.3.4. Solution design

One of the key requirements of the Solution was to add minimum overhead and changes to the pre-existing pseudonymisation and value-added service provision processes at the MNO, which is sketched in Figure 4 below:

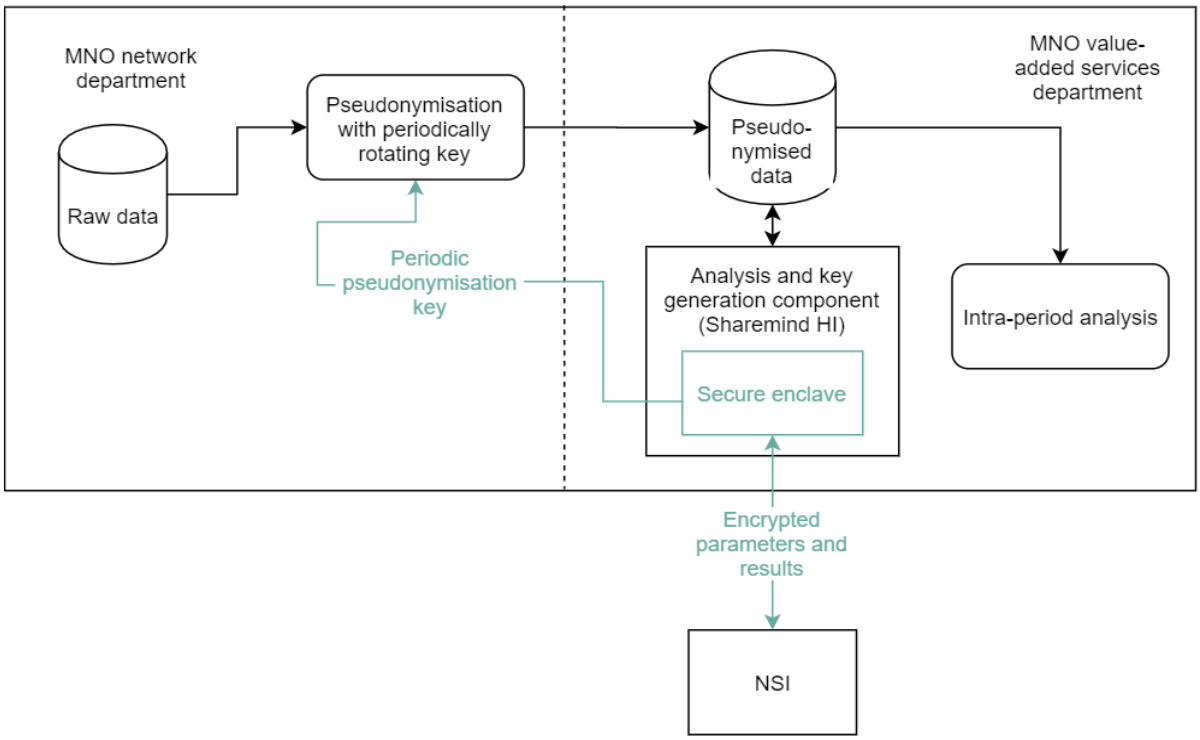
Figure 4 – The current process of sharing and analysing mobile location data[53]



[53] See: Solution Analysis. Figure 1: The current process of sharing and analysing mobile location data.

In order to meet the requirement to maintain the current processes at the MNO as well as other requirements deriving from the reference scenario, Toy Methodology and proof-of-concept setting,[54] Cybernetica AS proposed to implement the Solution on the Sharemind HI platform[55] using the architecture depicted in Figure 5 below. Sharemind HI development platform relies on a Trusted Execution Environment ("**TEE**") technology. A TEE isolates security sensitive parts of an application from the rest of the system with the help of trusted hardware. The TEE technology used in Sharemind HI is Intel® Software Guard Extensions ("**SGX**"), which is available in modern Intel® processors.[56] In the Solution architecture, the TEE is located within the Analysis and key generation component of Sharemind HI – it is comprised of "enclaves" in Intel® SGX terms (hereinafter "**enclave**" or "**enclaves (TEE)**").

Figure 5 – Proposed secure architecture of sharing and analysing mobile location data[57]



According to the proposed architecture, the Solution has two main functions:

1) **to generate periodic keys for pseudonymising mobile location data** – when the Solution is active, raw mobile location data is pseudonymised by the MNO-ND at its premises using the periodic pseudonymisation key generated in the enclave. While the enclave is physically on a server processor located at the MNO-VAD, MNO-VAD does not have access to or control over it, other than allowing the set-up and (de-)activation of the Solution. This means that

[54] The specific technical requirements are provided in the Solution Architecture Document.
[55] The Sharemind HI platform is described in more detail in the other deliverables of the Project. – See: Solution Analysis. Chapter 3.
[56] See: Solution Analysis Document. Section 3.3.
[57] See: Solution Analysis Document. Figure 3: Proposed secure architecture of sharing and analysing mobile location data.

MNO-VAD, or any third party for that matter, is not able to access or otherwise make use of the pseudonymisation keys stored in the enclave.

2) **to perform data analysis tasks on pseudonymised mobile location data** – as the periodic pseudonymisation keys are stored in the Trusted Execution Environment (TEE), they can be used to reverse the pseudonymisation of mobile location data inside the Trusted Execution Environment (TEE) in order to carry out data analysis. Both the reversal of pseudonymisation and the analysis are carried out in Solution enclaves within the TEE, ensuring that neither MNO-ND nor MNO-VAD has access to or is otherwise able to make use of the mobile location data. Only once the analysis is complete, will the TEE release pre-agreed reports in encrypted form, which can be decrypted by the NSI and MNO-VAD. The reports contain aggregated data, which have already been subjected to SDCs inside the TEE before the release.

The only modification in these pre-existing processes (see Figure 4 above) concerns the periodic key generation process[58] – the currently used change-and-forget method will be replaced by an analogous mechanism offered by means of the Solution. This does not result in a change in the business process, *per se*, however, it introduces the possibility to add new functionalities on top of existing ones.

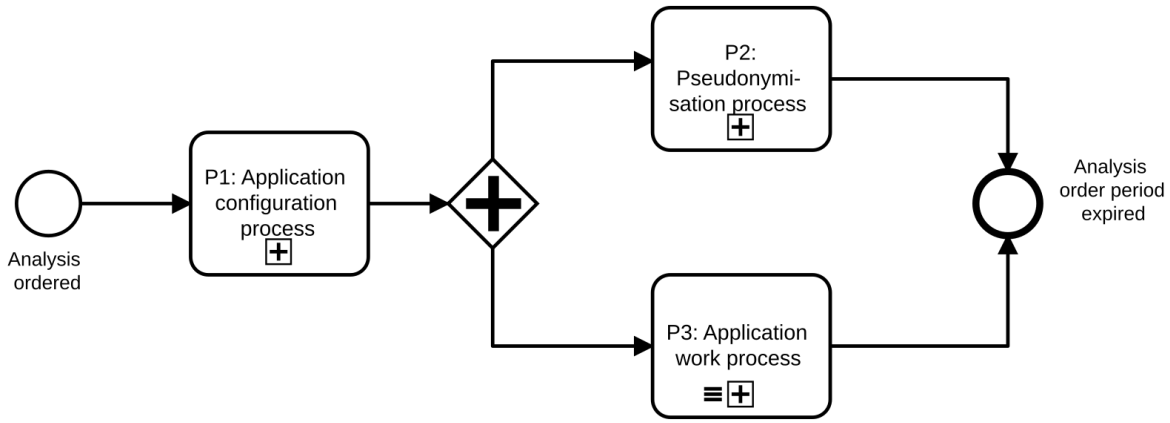
6.3.5. Process description

This section describes the processes that are carried out when running the Solution in the proof-of-concept and/or pilot stage. References to specific activities will be made based on the unique identification number assigned to such activities in the Solution Analysis.

The Solution is envisioned to perform the following processes (see Figure 6 below), each of which will be summarised in sections below:

a) Application Configuration Process (P1),
b) Pseudonymisation Process (P2),
c) Application Work Process (P3).

Figure 6 – Solution general process[59]

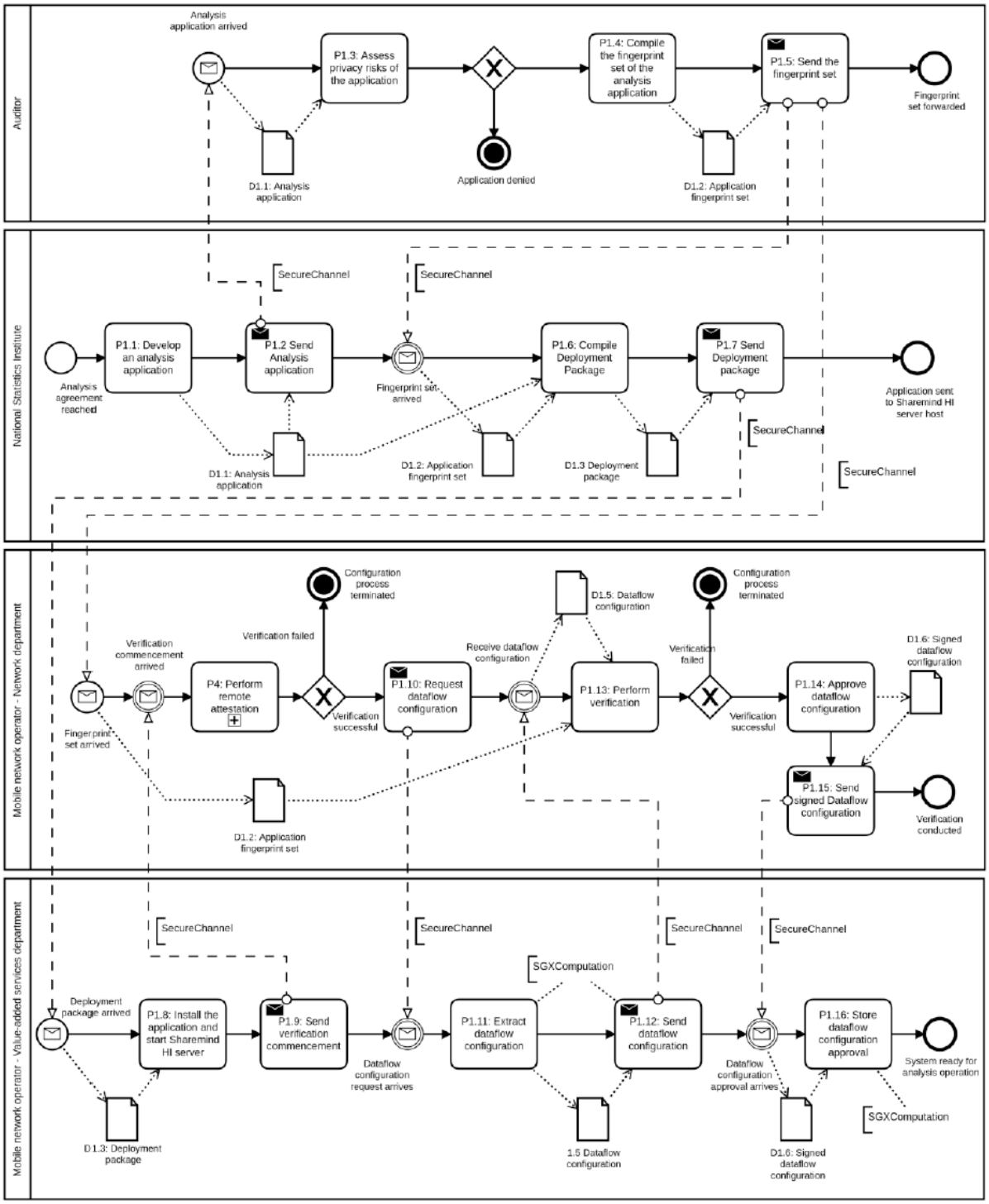


[58] See: Solution Analysis, Section 4.1.
[59] See: Solution Analysis. Figure 5: Solution general process.

i. Application Configuration Process (P1)

The details of the Application Configuration Process (P1) are provided in Figure 7 below.

Figure 7 - Application Configuration Process (P1)[60]



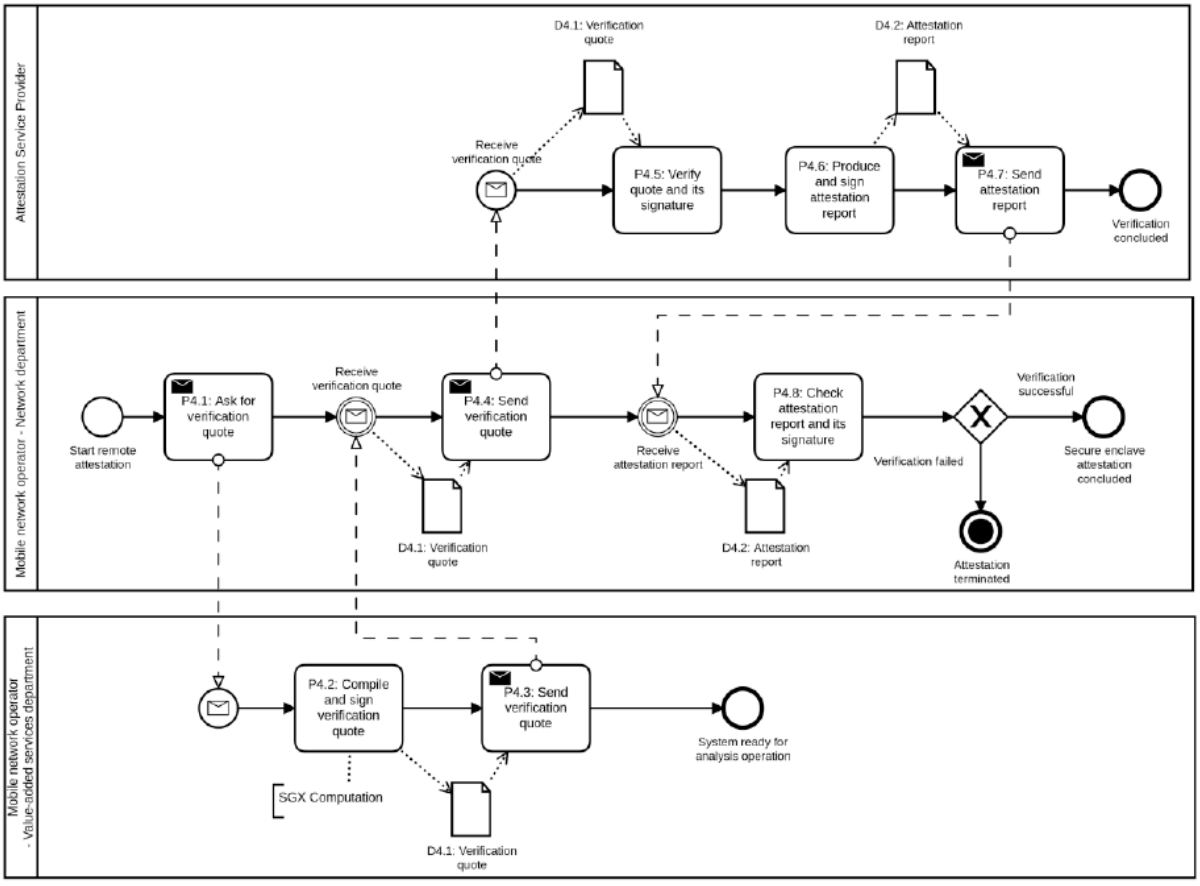
[60] See: Solution Analysis. Figure 6: Application configuration process.

The analysis of mobile location data will be carried out using an application, which is unique to each statistical analysis use case. The security and privacy risks are context-specific and require adaptation for the specific circumstances of each use case. The application developed for the Sample Use Case shall be hereafter referred to as "**Sample Use Case Application**".

In order to run the analysis in a privacy-preserving manner, the Sample Use Case Application needs to be developed, assessed for privacy risks, approved for implementation, set up, and attested, as further detailed in the Solution Analysis.[61]

The Remote Attestation Process (P4) is a separate sub-process carried out as part of all main processes, i.e. the Application Configuration Process (P1), the Pseudonymisation Process (P2) and the Application Work Process (P3) – its details are provided in Figure 8 below. Nevertheless, the Remote Attestation Process (P4) is excluded from the following analysis as it does not involve processing mobile location data.

Figure 8 - Remote Attestation Process (P4)[62]



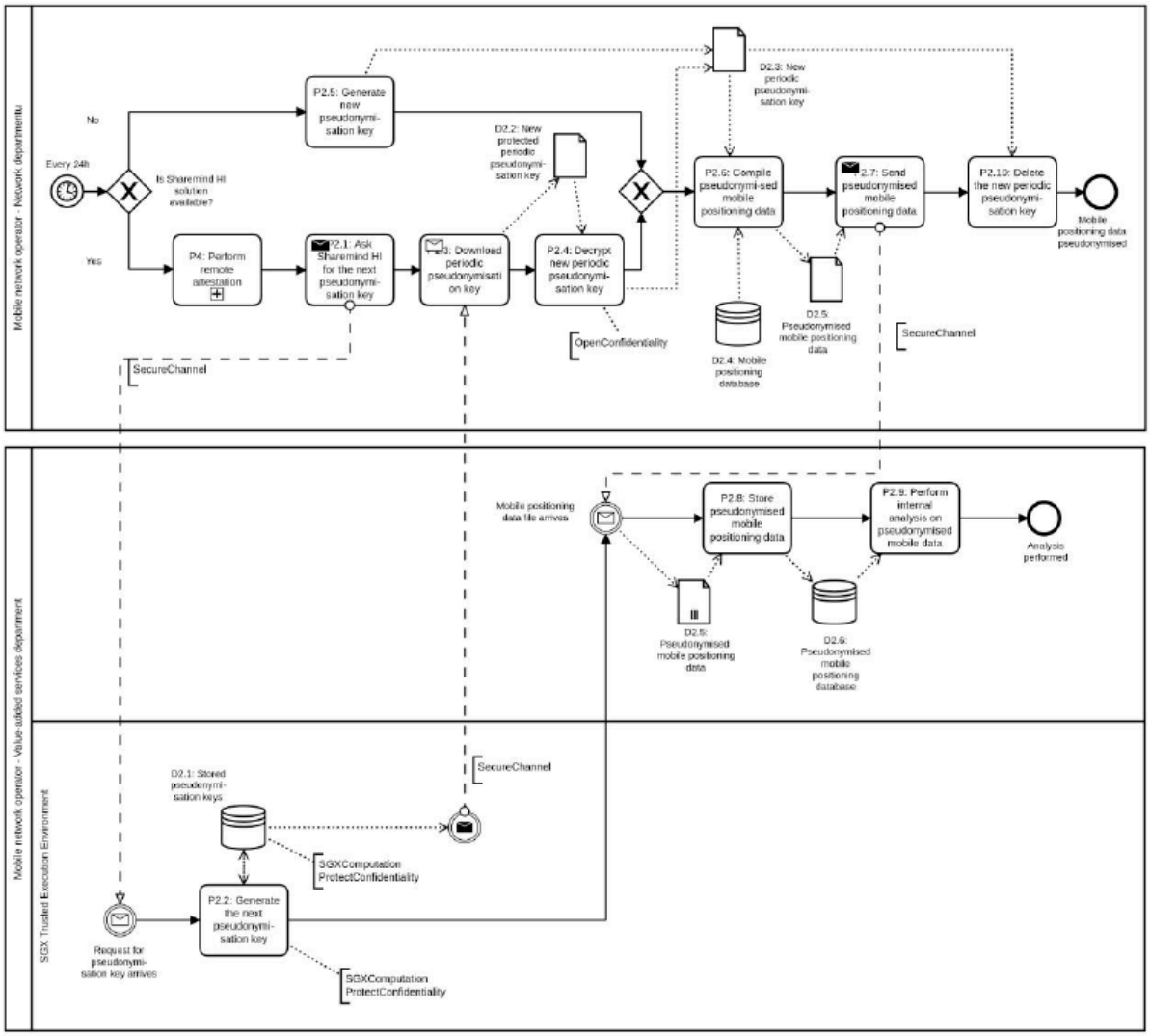
The Application Configuration Process (P1) and its sub-processes do not yet involve processing of mobile location data. For this reason, it is left aside for the purposes of the remaining Sample DPIA process.

[61] See: Solution Analysis. Section 5.2.
[62] See: Solution Analysis. Figure 9: Remote attestation process.

ii. Pseudonymisation Process (P2)

The details of the Pseudonymisation Process (P2) are provided in Figure 9 below.

Figure 9 - Pseudonymisation Process (P3)[63]



The Pseudonymisation Process (P2) involves processing of mobile location data. If the Application Configuration Process (P1) has been successfully completed, then:

1) the Solution will be attested (P4),
2) the Solution will be asked for a new periodic pseudonymisation key every 24h (P2.1),
3) the Solution will generate and store a new periodic pseudonymisation key for every 24h (P2.2),
4) the new periodic pseudonymisation key is downloaded (P2.3), decrypted (P2.4) and used to compile pseudonyms for the International Mobile Subscriber Identity ("**IMSI**") values (P2.6),

[63] See: Solution Analysis. Figure 7. MNO data pseudonymisation process.

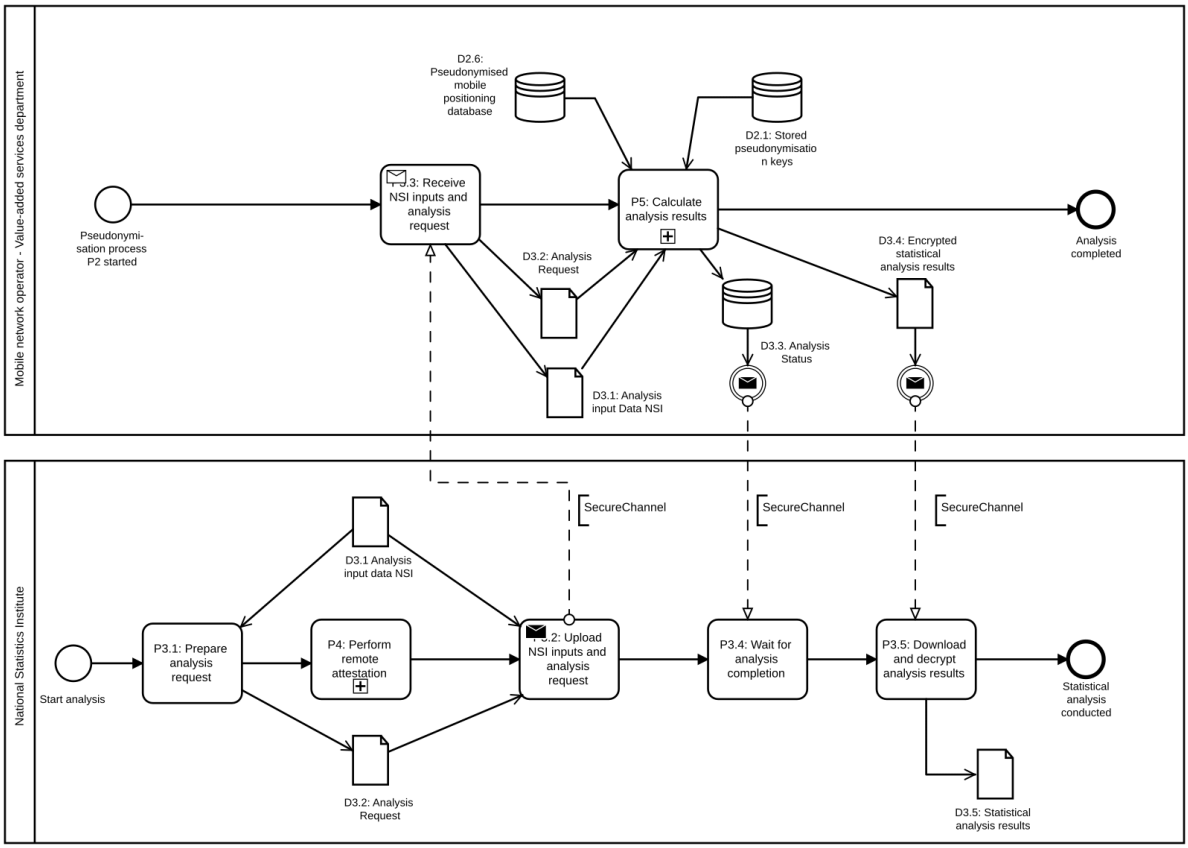
5) the new IMSI pseudonyms together with associated mobile location data for the current 24h cycle are sent (P2.7) and stored (P2.8),
6) intra-period analysis can be conducted using the pseudonymised mobile location data in accordance with the pre-existing value-added service provision process (P2.9),
7) the new periodic pseudonymisation key is deleted (P2.10).

If the Solution is unavailable, the new pseudonymisation keys will be generated using the pre-existing change-and-forget method (P2.5 and P2.10), in order not to create any friction in the pre-existing pseudonymisation and value-added service provision processes at the MNO.

iii. Application Work Process (P3)

The details of the Application Work Process (P3) are provided in Figure 10 below.

Figure 10 - Application Work Process (P3)[64]



The Application Work Process (P3) involves processing of mobile location data. It is dependent on the Pseudonymisation Process (P2), which provides the required input data – the former will run only if the latter is active and the Solution is available. If so, then:

1) preparations are made for requesting a statistical analysis (P3.1),
2) the Solution is attested (P4),
3) the statistical analysis is requested for (P3.2),

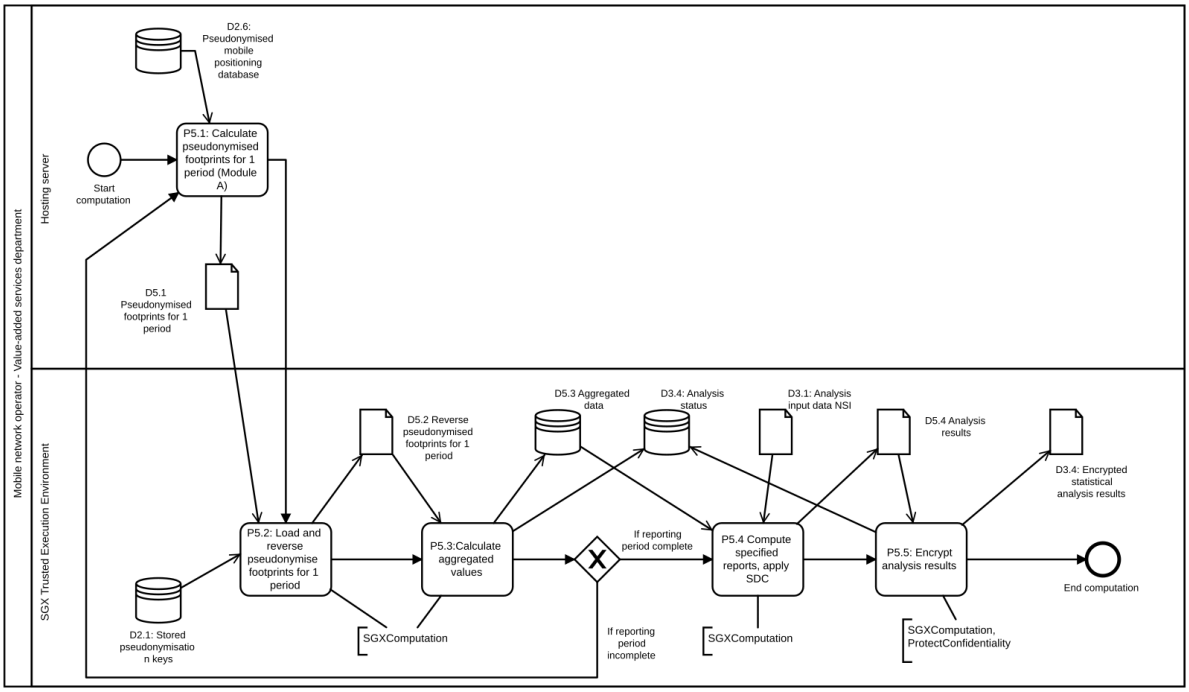
[64] See: Solution Analysis. Figure 8. Application work process.

4) the statistical analysis request is received (P3.3),
5) the statistical analysis is carried out (P5),
6) updates are received until completion of the statistical analysis (P3.4),
7) the results of the statistical analysis are downloaded and decrypted (P3.5).

If the Solution is unavailable, new pseudonymisation keys will not be generated using the Solution. In such case, the process will revert back to the pre-existing method that the MNO currently applies for providing value-added services. If so, there will be a loss of mobile location data available for the statistical analysis by means of the Solution.

As a sub-process of the Application Work Process (P3), the statistical analysis implements the Sample Use Case. The details of the respective Use Case Process P5) are provided in Figure

Figure 11 - Use Case Process (P5)[65]



The statistical analysis sub-process is conducted in two separate environments on the premises of the MNO:

1) **Module A** – the Intra-Period Footprint and Footprint Score are calculated as part of the pre-processing stage of the Toy Methodology outside the Solution (see Section 6.3.2.i.a) above).
2) **enclaves (TEE)** – the rest of the calculations as specified in the Toy Methodology are carried out as part of the Solution (see Section 6.3.2.i.b).

All the pre-approved calculations described in the Sample Use Case are run automatically in the enclaves (TEE). The enclaves (TEE) apply cryptographic techniques, which enable it to securely link mobile location data records over several

[65] See: Solution Analysis. Figure 10. Use case process.

pseudonymisation periods, while none of the stakeholders have access to or visibility of the pseudonymized mobile location data in the enclaves (TEE) during processing.

6.3.6. Data description

i. Input data

This section describes the data that is used as input when running the Solution in a proof-of-concept stage, except that input data will be real-world mobile location data instead of synthetic mobile location data. References to specific data elements will be made based on the name and unique identification number assigned to such elements in the Solution Analysis.

a) NSI input

NSI provides MNO with inputs to carry out the relevant processing steps of the Sample Use Case:

- a predefined algorithm for computing the Footprint Score (public information), the estimation of the load of movement between a city and its commuting areas, the estimation of the load of movement between a city and its commuting areas, the number of Observed Total Users and Adjusted Total Users.
- the Reference Areas (public information),
- the thresholds for determining most visited grid cells per mobile device (public information),
- the Resident Count (secret input), this is an optional data element that is used solely for calibration; if absent, no calibration is conducted; if present, such data element is secret;
- the threshold values for SDC, the SDC methods and the associated code.

b) MNO input

MNO makes available the pseudonymised mobile location data for pre-processing in Module A (outside the Solution's enclave), which includes:

- pseudonymised IMSI;
- timestamp;
- position.

During pre-processing in the Module A, the pseudonymised mobile location data is further temporally summarised, resulting in:

- pseudonymised IMSI;
- daily sub-period, as a result of temporal summarisation in Module A;
- position with 1km^2 grid cell precision, after spatial coarsening in Module A.

ii. Output data

The following reports and aggregate results are made available as output of the Solution:

1) Fingerprint Report,
2) Population Density Report,

3) FUF Report,
4) Highly Nomadic Users,
5) Observed Total Users,
6) Adjusted Total Users.

iii. Personal data description per process

According to the CNIL guide on PIA methodology,[66] the study of the context of the processing has to define and describe in detail the personal data concerned.

The Sample DPIA is focused on evaluating the impact of introducing mobile location data as a new source for computing the territorial dimension in official statistics by means of the Solution (see Section 6.3 above). Therefore, this document is dedicated to analysing the privacy-preserving processing of pseudonymous mobile location data by means of the Solution, as opposed to other types of personal data which may be provided as input or produced as output in the supporting processes of the Solution (e.g. user credentials and activity logs of the authorised users of the Solution). Such other types of personal data will be the subject of future legal analysis in the next iterations of the DPIA.

This document does not cover non-personal data processed in the Sample Use Case by means of the Solution. All non-personal data elements that were identified in other deliverables produced under the Agreement[67] have been excluded from the analysis contained herein. This includes, *inter alia*, the results of the collective aggregation step in Module D in the Trusted Execution Environment (see Figure 3 above), whereby the link between an individual mobile device and its location data is eliminated by deleting the underlying reverse pseudonymised mobile location data and the respective 24h pseudonymisation keys.

Although not part of the Solution, Module A also conducts personal data processing as part of the pre-processing step of the individual integration stage (see Figure 3 above). Since the results of the pre-processing in Module A are used as input for the Solution, it was considered necessary to include this step in the following analysis, in order to cover different processing operations at the macro level and better distinguish between processing with traditional technologies (Module A) and processing with privacy-enhancing technologies (the Solution).

The following tables in this section summarise the different types of mobile location data that are received by the relevant stakeholders (NSI and MNO) for further processing in the context of the Sample Use Case. Some of these personal data are initially collected by the MNO for purposes related to delivering telecommunications services and/or in support of network operation ("**Primary Processing**") and later re-used for the purposes of producing official statistics ("**Secondary Processing**"). The Sample DPIA is focused on the Secondary Processing, although Primary Processing is also covered to the extent it affects the Secondary Processing.

---

[66] CNIL. Privacy Impact Assessment (PIA). Methodology. February 2018 edition. – In the Internet: https://www.cnil.fr/sites/default/files/atoms/files/cnil-pia-1-en-methodology.pdf (24.08.2021).

[67] The excluded data elements are Analysis input data NSI (D3.1), Analysis request (D3.2), Aggregated data (D5.3), Encrypted statistical analysis results (D3.4) and Analysis results (D3.5, D5.4), as described in the document „ESTAT 2019.0232 Solution analysis".

None of the tables include personal data which is generated in the Primary Processing phase, but not later re-used in the Secondary Processing phase. If the re-use of data elements in the Secondary Processing phase includes extracts of data elements generated in the Primary Processing, then only the entire data element as a whole is referred to, in order to avoid confusion regarding potential double processing.[68]

All data elements described in the tables refer to a numeric ID assigned to such data element in the "ESTAT 2019.0232 Solution analysis document"[69] (in the format "Dn.n"). The same ID can be used to trace the relevant data element in the figures illustrating the process description in Section 6.3.5 above (see Figure 6 – Figure 11 above), as well as throughout the rest of the deliverables produced under the Agreement.

The non-identifiable state of the pseudonymous mobile location data within the Solution is marked with the colour green in the following tables. The importance of highlighting such data is explained in Section 8.2 below, which provides the legal reasoning for treating this data as anonymous.

For the sake of better readability, the types of personal data processed have been divided into two tables:

1) Table 1 below represents the types of personal data relevant in the Primary Processing phase,
2) Table 2 below represents the types of personal data relevant in the Secondary Processing phase.

Table 1 – Description of data in Primary Processing

Data types	Recipients	Description of relevant data elements and processing activities
Raw mobile location data (D2.4)	MNO-ND	D2.4 – a database containing the following information: IMSI (International Mobile Subscriber Identity), timestamp and position. The MNO-ND creates the data element D2.4 as part of its usual business operations and maintains control over it.
24h pseudonymisation keys (D2.1, D2.2, D2.3)	Trusted Execution Environment (TEE)	D2.1 – a database composed of pseudonymisation keys, where each key is an independent random value valid for one period only (24h in the Sample Use Case). D2.2 – encrypted version of a pseudonymisation key extracted from the data element D2.1 for one period only (24h in the Sample Use Case).

[68] This concerns data elements D2.2 and D.2.3, which are extracts of data element D2.1 (encrypted and decrypted version, respectively).
[69] See: Solution Analysis.

		The TEE generates a new pseudonymisation key upon request from the MNO-ND as per the change-and-forget method and stores it in data element D2.1. The TEE encrypts the pseudonymisation key as data element D2.2 and sends it to the MNO-ND. The TEE protects the generated keys.
	MNO-ND	D2.2 – encrypted version of a pseudonymisation key extracted from the data element D2.1 for one period only (24h in the Sample Use Case). D2.3 – decrypted version of data element D2.2. The MNO-ND receives the data element D2.2 from the TEE, obtains data element D2.3 as a result of decrypting data element D2.2 and uses data element D2.3 in accordance with the change-and-forget method in its usual business operations.
Pseudonymised mobile location data (D2.5, D2.6)	MNO-ND	D2.5 – the pseudonymised version of an extract from the data element D2.4. The MNO-ND pseudonymises extracts of the data element D2.4, using the data element D2.3, and sends them to MNO-VAD as data element D2.5.
	MNO-VAD	D2.5 – the pseudonymised version of an extract from the data element D2.4. D2.6 – a database composed of data elements D2.5 over several periods. The MNO-VAD receives data elements D2.5 from the MNO-ND and stores them for further processing in data element D2.6 as part of its usual business operations.

Table 2 – Description of data in Secondary Processing

Data types	Recipients	Description of relevant data elements and processing activities
Pseudonymised mobile location data (D2.6)	MNO-VAD	D2.6 – a database composed of data elements D2.5 over several periods. The MNO-VAD sends data element D2.6 to the Module A.
Temporally summarised pseudonymised mobile location data (D5.1)	MNO-VAD (Module A)	D5.1 – a temporally summarised version of the data element 2.6. The Module A temporally summarises the data element D2.6, receives data element D5.1 as the output and sends it to the TEE.

	Trusted Execution Environment (TEE)	D5.1 – a temporally summarised version of the data element D2.6. The TEE receives the data element D5.1 from the Module A, stores and protects it for further processing by means of the TEE.
24h pseudonymisation keys (D2.1)	Trusted Execution Environment (TEE)	D2.1 – a database composed of pseudonymisation keys, where each key is an independent random value valid for one period only (24h in the Sample Use Case). The TEE uses the data element D2.1 to reverse pseudonymise data element D5.1. The TEE stores and protects the data element D2.1 for further processing by means of the TEE.
Temporally summarised reverse pseudonymised mobile location data (D5.2)	Trusted Execution Environment (TEE)	D5.2 – a reverse pseudonymised version of data element D5.1. The TEE reverse pseudonymises data element D5.1, receives data element D5.2 as the output and calculates aggregated values from it (the latter of the two can no longer be linked to individual Subscribers). The TEE stores and protects the data element D5.2 for further processing by means of the TEE.

6.4. Stakeholder description

According to the DPIA Methodology, the next step in the preparation stage of a DPIA after describing the target of evaluation is to identify actors involved and persons concerned. "Aside from organizations and persons participating in the development or implementation (and thereby potential attackers), all persons affected by the use should be involved, such as

- *the manufacturer of the test object,*
- *operators e.g. as processors (data centers, internet service providers),*
- *the controller employees,*
- *the persons concerned in their respective roles as citizens, patients, customers, employees, etc.,*
- *third parties who take note of personal data, either by chance (persons randomly present) or by intent (security services)."*[70]

This section describes the stakeholders who are involved in running the Solution in the proof-of-concept and/or pilot project stage.

i. NSI

NSI acts in the general role of the initiator and direct beneficiary of further processing pseudonymous mobile location data by means of the Solution for producing official statistics. In this role, the NSI:

[70] *Op.cit.*, F. Bieker et al. A Process for Data Protection Impact Assessment, 2016, pp 28-29.

- selects the statistical analysis use case, along with appropriate statistical methodologies, suitable for implementing in a real-world scenario by means of the Solution,
- customises the Solution for the purposes of the selected statistical analysis use case,
- contacts the MNO for making the mobile location data of its Subscribers available for statistical analysis by means of the Solution,
- provides the MNO with the customised Solution,
- activates the Solution,
- provides the MNO with the statistical analysis report request,
- provides the necessary public and private inputs for running the computations on the mobile location data and calibrating the results by means of the Solution, and
- receives the outputs from the Solution.

ii. MNO

MNO acts in the general role of a data source and service provider for the NSI. In this role, the MNO:

- agrees to make available the mobile location data of its Subscribers for statistical analysis by means of the Solution,
- configures the Solution,
- pseudonymises the mobile location data of its Subscribers by means of the Solution,
- makes the pseudonymised mobile location data of its Subscribers available for statistical analysis by means of the Solution,
- fulfils the statistical analysis report request received from the NSI by means of the Solution,
- implements the necessary public and private inputs for running the computations on the mobile location data of its Subscribers and calibrating the results by means of the Solution,
- produces the outputs by means of the Solution and receives a copy thereof.

It is presumed that there is an internal separation of functions between two different departments within the MNO – namely, the Network Department of the MNO ("**MNO-ND**") and the Value-Added Services Department of the MNO ("**MNO-VAD**") – enforced by means of different access rights and technical protection measures for each department. MNO-ND is responsible for the pseudonymisation of mobile location data, whereas all other activities are carried out by the MNO-VAD.

iii. Eurostat

Eurostat acts in the general role of a facilitator and promoter of processing mobile location data by means of the Solution for producing official statistics. In this role, Eurostat developed the reference scenario. The Toy Methodology and the requirements for the proof-of-concept setting were developed by Eurostat in consultation with Cybernetica AS.

iv. Cybernetica AS

Cybernetica AS acts in the general role of the developer and provider of the Sharemind HI platform underlying the Solution. In this role, Cybernetica AS developed the Solution based on the reference scenario provided by Eurostat. Cybernetica consulted Eurostat in developing the Toy Methodology and the requirements for the proof-of-concept setting.

In the proof-of-concept and/or pilot project stage, Cybernetica AS also acts as a proxy to the Attestation Service for Intel® SGX (IAS).

v. Intel, Inc.

Intel, Inc. will act as the Attestation Service Provider in the pilot project and/or production stage.

vi. DPA

The DPA acts in the general role of a national data protection supervisor and regulator in the relevant Member State where, simultaneously:

1) the NSI is located and authorised to act as the producer of official statistics, and
2) Subscribers to the MNO's mobile network can be located during the period for which the NSI has sent a statistical analysis report request to the MNO.

Throughout the different stages of adopting the Solution in practice (from proof-of-concept to production), the DPA is expected to authorise the activation of the Solution for further processing pseudonymous mobile location data for the purposes of producing official statistics. It does so as a result of an assessment of privacy risks of the application.

vii. Subscriber

Subscribers are the data subjects whose mobile location data is provided for analysis by means of the Solution for the purposes of producing official statistics.

7. Legal requirements relevant for the Sample DPIA

According to the DPIA Methodology, the next step in the preparation stage of a DPIA after identifying the actors involved and persons concerned is to present the relevant applicable laws (e.g. national laws of Member States which specify the GDPR in certain legal aspects, as well as sector specific legislation, inter alia, for the areas of telecommunications, social security, rules on professional secrecy or the protection of minors). "However, as a DPIA deals with processes and technical operations, these rules are only of concern if they are implemented directly in the process."[71] For the purposes of the Project, the national laws of a specific Member State were excluded from the analysis due to the fact that the Member State, the domain of official statistics and any relevant production process are unspecified in the Sample Use Case.

7.1. Data protection law

7.1.1. Overview

Data protection law is largely harmonized in the EU, thanks to the directly applicable GDPR, which took effect on 25.05.2018. The GDPR lays down general rules relating to the protection of natural persons with regard to the processing of personal data and rules relating to the free movement of personal data.[72]

In parallel with the GDPR, EDPR lays down rules relating to the protection of natural persons with regard to the processing of personal data by the Union institutions and bodies and rules relating to the free movement of personal data between them or to other recipients established in the Union.[73] Whenever the provisions of EDPR follow the same principles as the provisions of GDPR, those two sets of provisions should, under the case law of the Court of Justice of the European Union ("**ECJ**"), be interpreted homogeneously, in particular because the scheme of the EDPR should be understood as equivalent to the scheme of the GDPR.[74] For the purposes of the Sample DPIA, it is assumed that Eurostat – a Union institution – will not be participating as a direct stakeholder in any of the potential statistical analysis use cases to be implemented by means of the Solution in the future, i.e. it will not be taking the role of an NSI. Therefore, the EDPR is not analysed separately and the rest of the Sample DPIA is focused on the analysis of GDPR and DPD.

Many of the principles in the data protection legislation were established already in DPD, the predecessor of the GDPR. However, the GDPR did introduce some

[71] *Op. cit.*, F. Bieker et al. A Process for Data Protection Impact Assessment. 2016, p 29.
[72] GDPR Art 1(1), EDPR Rec 3.
[73] EDPR Art 1.
[74] EDPR Rec 5, referring to ECJ judgment of 9 March 2010, European Commission v Federal Republic of Germany, Case C-518/07, ECLI:EU:C:2010:125 paragraph 28. – See: European Data Protection Supervisor. The EDPS Strategy 2020-2024. Shaping a Safer Digital Future, footnote 1, p 6. – Internet: https://edps.europa.eu/sites/default/files/publication/20-06-30_edps_shaping_safer_digital_future_en.pdf (07.04.2021).

novelties compared to the DPD. The most relevant changes for the purposes of this legal analysis are outlined in the following section.

i. Special regime for statistics

One of the main changes as a result of the GDPR is that the rules on processing personal data primarily and directly follow the EU law, while regulations under national law are merely supplementary, to the extent that GDPR allows it.[75]

A direct result of this change, as recently confirmed by the EDPS in "A Preliminary Opinion on data protection and scientific research"[76] ("**EDPS Preliminary Opinion on Research**"), was the introduction of a special regime for scientific research under the GDPR. According to EDPS, the special regime is composed of specific derogations from certain controller obligations plus a specific provision (GDPR Art 89) requiring appropriate safeguards.[77] At the same time, EDPS admits that the full extent of this special regime is not precisely delineated because GDPR affords Member States with restricted flexibility to provide for derogations from the data protection rights, if there is no harmonised EU law for scientific research (e.g. clinical trials). The EDPS reiterates that "the special regime cannot be applied in such a way that the essence of the right to data protection is emptied out, and this includes data subject rights, appropriate organisational and technical measures against accidental or unlawful destruction, loss or alteration, and the supervision of an independent authority".[78]

What seems to be clear is that the use of most exceptions to certain requirements of the GDPR foreseen in case of processing personal data for scientific research is made conditional on having in place additional safeguards as required by GDPR Art 89(1).[79] In its recent Document on response to the request from the European Commission for clarifications on the consistent application of the GDPR, focusing on health research ("**EDPB Document on Health Research**"), the EDPS has admitted the present lack of specification on what could or should be considered adequate safeguards under GDPR Art 89(1).[80] This "can be considered a serious impediment for the proper use of the exceptions foreseen in the GDPR for processing personal data for scientific research purposes."[81] Nevertheless, the EDPB argues, "without

---

[75] Eds. T. Runge, M. Bug, K. Schaar. Data Protection Guide. 2nd fully revised edition. RatSWD Output 8 (6). Rat für Sozial- und Wirtshaftsdaten (German Data Forum), Berlin, October 2020, p 8, 14. – In the internet: https://www.konsortswd.de/en/latest/publication/data-protection-guide-2nd-edition/ (25.01.2021).
[76] European Data Protection Supervisor. A Preliminary Opinion on data protection and scientific research. 6 January 2020, p 2, p 5, 19. – Internet: https://edps.europa.eu/data-protection/our-work/publications/opinions/preliminary-opinion-data-protection-and-scientific_en (08.04.2021).
[77] *Ibid.*, p 16, p 19.
[78] *Ibid.*, p 19.
[79] European Data Protection Board. EDPB Document on response to the request from the European Commission for clarifications on the consistent application of the GDPR, focusing on health research. Adopted on 2 February 2021, p 12, sec 55. – Internet: https://edpb.europa.eu/our-work-tools/our-documents/other-guidance/edpb-document-response-request-european-commission_en (07.05.2021).
[80] *Ibid.,* p 12, sec 53.
[81] *Ibid.*, p 12, sec 55.

such – yet to be clarified – additional safeguards the use of such research exceptions would not be legitimate."[82]

To conclude, scientific research area finds itself in a paradoxical situation – on the one hand, they are entitled to a more lenient application of GDPR on the condition the appropriate safeguards are implemented; on the other hand, the criteria for determining the appropriateness of the safeguards are not settled in practice nor clarified by relevant data protection authorities.

The norms referred to by EDPS in support of its argumentation for a privileged position of scientific research and the relating condition of appropriate safeguards in the GDPR also expressly mention, *inter alia*, the processing for statistical purposes. The regulation of data processing for scientific research and statistical purposes was regulated in conjunction also under the DPD.[83] For these reasons, it is reasonable to expect that if GDPR created a special regime for scientific research, it did the same also for processing for statistical purposes. Furthermore, if the criteria for appropriate safeguards is unsettled for scientific research under the GDPR, the situation is the same also for statistical purposes. Therefore, the special regimes for both scientific research and statistical purposes are currently in a state of flux. For these reasons, interpretations of GDPR rules in the context of scientific research can be extended also to processing for statistical purposes. However, drawing any analogies should be done carefully as there are principal and practical differences between scientific research and official statistics, starting with the different public interest purposes and ending with the complexity of domain-specific regulation.

ii. Additional rules on deidentified data

The GDPR explicitly mentions pseudonymisation, where as the DPD remained silent on the matter. GDPR draws an explicit line between:

1) **pseudonymous data** – data which has undergone pseudonymization and classifies as personal data (information on an identifiable natural person);
2) **anonymous data** – information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable.[84]

Compared to DPD, the GDPR also introduced new rules on processing which does not require identification, embodied in GDPR Art 11. If the purposes for which a controller processes personal data do not or do no longer require the identification of a data subject by the controller, the controller shall not be obliged to maintain, acquire or process additional information in order to identify the data subject for the sole purpose of complying with GDPR.[85] Additionally, if the controller is able to demonstrate that it is not in a position to identify the data subject, the controller shall inform the data subject accordingly, if possible.[86] In such cases, certain rights of the data subject (right of access, right of rectification, right of erasure, right to restriction

[82] *Ibid.*
[83] DPD Rec 29, 34, 40; DPD Art 6(1)(b) and (e), Art 11(2).
[84] GDPR Rec 26.
[85] GDPR Art 11(1).
[86] GDPR Art 11(2).

of processing, notification obligation regarding rectification or erasure of personal data or restriction of processing, right to data portability)[87] shall not apply except where the data subject, for the purpose of exercising those rights, provides additional information enabling his or her identification.[88] This does not release the controller from the obligations related to other rights of the data subject, such as:

1) transparency, information and access to personal data[89],
2) right to object,[90]
3) right not to be subject to automated individual decision-making, including profiling[91].

7.1.2. Obligation to carry out a data protection impact assessment

Controllers have a general obligation to implement appropriate technical and organisational measures to ensure and to be able to demonstrate that processing is performed in accordance with the GDPR.[92] When performing this obligation, controllers have to take into account the nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for the rights and freedoms of natural persons.[93] This means that controllers must continuously assess and manage the risks for the rights and freedoms of natural persons created by their processing activities.[94]

Similarly, both controllers and processors have a general obligation to implement appropriate technical and organisational measures to ensure a level of security appropriate to the risk, including *inter alia* as appropriate:

(a) the pseudonymisation and encryption of personal data;
(b) the ability to ensure the ongoing confidentiality, integrity, availability and resilience of processing systems and services;
(c) the ability to restore the availability and access to personal data in a timely manner in the event of a physical or technical incident;
(d) a process for regularly testing, assessing and evaluating the effectiveness of technical and organisational measures for ensuring the security of the processing.[95]

The obligation to carry out a DPIA arises only where there is a likelihood of a high risk to the rights and freedoms of natural persons due to the nature, scope, context and purposes of the data processing.[96] High risk processing includes profiling, large scale processing of special categories of data and systematic monitoring of a publicly accessible area on a large scale.[97] However, a DPIA may also be required in

[87] GDPR Art-s 15-20.
[88] GDPR Art 11(2).
[89] GDPR Art-s 12-14.
[90] GDPR Art 21.
[91] GDPR Art 22.
[92] GDPR Art 24(1).
[93] *Ibid*.
[94] *Op. cit*. Article 29 Data Protection Working Party. Guidelines on Data Protection Impact Assessment (DPIA), 2017, p 6.
[95] GDPR Art 32(1).
[96] GDPR Art 35(1).
[97] GDPR Art 35(3).

cases which fulfil any of the following nine criteria – the more criteria are met by the processing, the more likely it is to present a high risk to the rights and freedoms of data subjects, regardless of the measures which the controller envisages to adopt[98]:

1) evaluation or scoring, including profiling and predicting;
2) automated decision-making with legal or similar significant effect – processing that aims at taking decisions on data subjects;
3) systematic monitoring – processing used to observe, monitor or control data subjects, including data collected through networks;
4) sensitive data or data of highly personal nature;
5) data processed on a large scale, considering the number of data subjects concerned, the volume of the data and/or the range of different data items being processed, the duration or permanence of the data processing activity, the geographical extent of the processing activity;
6) matching or combining datasets;
7) data concerning vulnerable data subjects;
8) innovative use or applying new technological or organizational solutions.[99]

Once a DPIA has been carried out, GDPR Art 35(11) further obliges the controller, upon necessity, to carry out a review to assess if processing is performed in accordance with the DPIA at least when there is a change of the risk represented by processing operations.

If it is not clear whether a DPIA is required, the WP29 recommends that a DPIA is carried out nonetheless as it is a useful tool to help controllers comply with data protection law.[100]

7.1.3. Obligation to consult the supervisory authority

There are three situations for consulting the supervisory authority[101] when conducting a DPIA:[102]

1) **controller's obligation under the GDPR** – prior to processing in cases where a DPIA indicates that the processing would result in a high risk in the absence of measures taken by the controller to mitigate the risk.[103] For the purposes of such consultation, the controller shall provide the supervisory authority with:
    a. the respective responsibilities of the controller, joint controllers and processors involved in the processing, where applicable;
    b. the purposes and means of the intender processing;
    c. the measures and safeguards provided to protect the rights and freedoms of data subjects pursuant to the GDPR;

[98] *Op. cit*. Article 29 Data Protection Working Party. Guidelines on Data Protection Impact Assessment (DPIA), 2017, p 11.
[99] *Ibid.*, pp 9-11.
[100] *Ibid.*, p 8.
[101] According to GDPR Art 4(1)(21), supervisory authority means an independent public authority which is established by a Member State pursuant to GDPR Art 51.
[102] GDPR Art 36.
[103] GDPR Art 36(1).

d. the contact details of the DPO, where applicable;
e. the DPIA[104];
f. any other information requested by the supervisory authority.[105]

2) **obligation of Member States under the GDPR** – during the preparation of a proposal for a legislative measure to be adopted by a national parliament, or of a regulatory measure based on such a legislative measure, which relates to processing;[106]

3) **right of Member States under the GDPR** – Member State law may require controllers to consult with, and obtain prior authorisation from, the supervisory authority in relation to processing by a controller for the performance of a task carried out by the controller in the public interest, including processing in relation to social protection and public health.[107]

7.1.4. Processing personal data for statistical purposes

i. Legal definition of statistical purposes

The GDPR applies whenever personal data is processed for statistical purposes. Statistical purposes mean any operation of collection and the processing of personal data necessary for statistical surveys or for the production of statistical results. Further, the statistical purpose implies that the result of processing for statistical purposes is not personal data, but aggregate data, and that this result or the personal data are not used in support of measures or decisions regarding any particular natural person.[108]

The EU data protection law distinguishes official statistics as a subset of statistics in general.[109] Official statistics are divided into official European and official national statistics. European statistics should be developed, produced and disseminated in accordance with the statistical principles as set out in Art 338(2) of the TFEU and further specifications on statistical confidentiality in the RES.[110] National statistics should also comply with Member State law.[111]

ii. Purpose limitation

a) Introduction

[104] The WP29 considers this as a reference to a DPIA report. – See: *Op. cit*. Article 29 Data Protection Working Party. Guidelines on Data Protection Impact Assessment (DPIA), 2017, p 20.
[105] GDPR Art 36(3).
[106] GDPR Art 36(4)
[107] GDPR Art 36(5).
[108] GDPR Rec 162.
[109] GDPR Rec 163, EDPR Rec 85.
[110] *Ibid*.
[111] GDPR Rec 163.

Purpose limitation is one of the key data protection principles. It requires that personal data shall be collected for "specified, explicit and legitimate purposes" and "not further processed in a manner that is incompatible with those purposes".[112]

In 2013, in light of the discussion on the precise meaning of and exceptions to the principle of purpose limitation, the WP29 adopted an opinion on purpose limitation to clarify the exact scope and function of this important principle ("**WP29 Opinion on Purpose Limitation**").[113] This work was largely motivated by a legislative proposal in the process which ultimately lead to the adoption the GDPR. According to this proposal, data processing for incompatible use was allowed provided a new legal ground is available (except legitimate interests of the controller), disconnecting the new data processing operation from the original purpose.[114] In the WP29 Opinion on Purpose Limitation, the WP29 provided arguments against such proposal and offered alternative wording, which was later used as basis for drafting the compatibility assessment requirement in GDPR Art 6(4).[115]

In the introduction of its Opinion on Purpose Limitation, the WP29 clarified that "[s]pecification of purpose is an essential first step in applying data protection laws and designing data protection safeguards for any processing operation. […] The principle of purpose limitation is designed to establish the boundaries within which personal data collected for a given purpose may be processed and may be put to further use. […] Indeed, the principle of purpose limitation inhibits 'mission creep', which could otherwise give rise to the usage of the available personal data beyond the purposes for which they were initially collected."[116]

Next, the WP29 provided a brief history of the purpose limitation in its Opinion on Purpose Limitation, concluding that the principle is clearly established in Art 8(2) of the Charter with two separate and distinct requirements – purpose specification and legitimate basis for processing.[117] Based on this distinction, the WP29 developed two main building blocks of purpose limitation, in order to aid the analysis of the concept:

1) **purpose specification** – data are collected for certain aims, which are the *raison d'être* of the processing operations. It will determine the relevant data to be collected, retention periods, and all other key aspects of how personal data will be processed for the chosen purpose/s;[118]

---

[112] GDPR Art 5(1)(b).
[113] Article 29 Data Protection Working Party. Opinion 03/2013 on purpose limitation. Adopted on 2 April 2013. 00569/13/EN WP 203, p 11. – Internet: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf (14.04.2021). Note that the WP29 Opinion 03/2013 on purpose limitation has not been endorsed by the EDPB. See: European Data Protection Board. GDPR: Guidelines, Recommendations, Best Practices. Endorsement of GDPR WP29 Documents. – In the Internet: https://edpb.europa.eu/our-work-tools/general-guidance/gdpr-guidelines-recommendations-best-practices_en (21.01.2021).
[114] *Ibid*.
[115] European Data Protection Supervisor, A Preliminary Opinion on data protection and scientific research, 6 January 2020, p 22, footnote 128. – Internet: https://edps.europa.eu/sites/edp/files/publication/20-01-06_opinion_research_en.pdf (18.05.2021).
[116] *Op cit.*, Article 29 Data Protection Working Party. Opinion 03/2013 on purpose limitation, 2013, p 4.
[117] *Ibid.*, pp 6-11.
[118] *Ibid.*, pp 11-12.

2) **compatible use** – the prohibition of incompatible use sets a limitation on further use, requiring that a distinction be made between "compatible further use" and "incompatible further use".[119]

b) Purpose specification

In the context of integrated farm statistics, the EDPS has considered it useful to distinguish between two different phases of statistical data processing:[120]

a) **preparation phase** – the initial phase while re-linking the data is still possible and desired, in order to combine and enrich statistical data by linking various datasets;
b) **later phase** – statistical data has been prepared and the keys allowing linking the various datasets can be destroyed.

Preparation phase

During the preparation phase, in order to ensure statistical confidentiality, statistical institutions dissociate the input data gathered from surveys and other sources, i.e. they pseudonymise the input data and also ensure that other technical and organisational measures are in place to minimise the risk that the individuals can be re-identified.[121]

"This process usually includes key-coding the data and ensuring that the keys, that is, information to link the datasets to the individuals whom they relate to, are kept separately."[122] Typically, the keys are instrumental in relinking the survey data to additional datasets that are necessary for the production of official statistics. The keys are not destroyed right away, in order to allow statistical institutes to combine and enrich the survey data with data from other sources.[123]

Later phase

In the later phase, once re-linking is no longer necessary for the statistical purposes sought, the keys are usually destroyed. Destruction of keys is typically a minimum requirement for eliminating possibilities to directly link individuals to the datasets using the keys at the earlier possible time. However, additional measures may often also need to be taken to ensure statistical confidentiality.[124]

---

[119] *Ibid.*, pp 12-13.
[120] European Data Protection Supervisor. Opinion 10/2017. EDPS Opinion on safeguards and derogations Under Article 89 GDPR in the context of a proposal for a Regulation on integrated farm statistics. 20 November 2017, p 9. – Internet: https://edps.europa.eu/data-protection/our-work/publications/opinions/farm-statistics_en (14.04.2021). 9.
[121] *Op cit.*, European Data Protection Supervisor. EDPS Opinion on safeguards and derogations under Article 89 GDPR in the context of a proposal for a Regulation on integrated farm statistics, 2017, p 9.
[122] *Ibid.*
[123] *Ibid.*
[124] *Ibid.*, p 10.

Whether or not retaining all or some of the original input (raw) data is appropriate and compliant with the principle of data minimisation needs to be assessed case-by-case. This may include entire datasets stripped of direct identifiers such as names and addresses, which do not meet the "high standards required for anonymisation".[125] "If appropriate anonymisation techniques have been applied and these ensure that the aggregated statistical datasets no longer contain any personal data, […] the GDPR will no longer be applicable to these fully aggregated and anonymised datasets".[126]

To conclude, the EDPS has acknowledged that personal data used in a typical statistical analysis process can be considered as anonymised only at the very end of the process, after aggregating the relevant datasets and applying appropriate anonymisation techniques.

With regard go the original input (raw) data that may need to be kept for longer periods of time, even after the keys are destroyed, the EDPS has acknowledged the technical difficulties of re-linking the files and suggested GDPR Art 11 to be relevant in this situation. Namely, in cases where a controller is able to demonstrate that it is not in a position to identify the data subject – such as when the keys are destroyed and other technical and organisational measures are taken to ensure that the individuals cannot be identified – the rights of the data subjects under GDPR Art-s 15-20 shall not apply.[127]

c) Compatible use

If the processing concerns personal data which has been collected earlier (further processing) and is carried out for a new purpose, the controller must carry out a compatibility assessment, in order to ascertain whether processing for another purpose is compatible with the purpose for which the personal data were initially collected. [128] Such an assessment must take into account, but not be limited to:

(a) any link between the purposes for which the personal data have been collected and the purposes of the intended further processing;
(b) the context in which the personal data have been collected, in particular regarding the relationship between data subjects and the controller;
(c) the nature of the personal data, in particular whether special categories of personal data are processed, or whether personal data related to criminal convictions and offences are processed;
(d) the possible consequences of the intended further processing for data subjects;
(e) the existence of appropriate safeguards, which may include encryption or pseudonymisation.[129]

---

[125] *Ibid.*
[126] *Ibid.*
[127] *Op cit.*, European Data Protection Supervisor. EDPS Opinion on safeguards and derogations under Article 89 GDPR in the context of a proposal for a Regulation on integrated farm statistics, 2017, p 10.
[128] GDPR Art 6(4).
[129] *Ibid*.

The compatible use building block of purpose limitation has included the so-called presumption of compatibility in the EU data protection law since the DPD.[130] The WP29 acknowledged already in 2013 that it was not clear from the text of DPD Art 6(1)(b) alone whether the specific provision on presumption of compatibility "should be seen as an *exception* to the general prohibition of incompatible use in order to give a privileged position to 'historical, statistical or scientific purposes' or as a *specification* of the general rule, while not excluding that other cases could also be considered as 'not incompatible'".[131] It took the position in favour of the latter view, whereby "the specific provision could give rise to more general criteria for compatibility (e.g. potential impact on the data subject, and appropriate safeguards)", which in turn led to "a more prominent role for different kinds of safeguards, including technical and organisational measures to ensure functional separation, such as full or partial anonymisation, pseudonymisation, aggregation of data, and privacy-enhancing technologies".[132]

According to the presumption of compatibility under the GDPR, the further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with GDPR Art 89(1), not be considered to be incompatible with the initial purposes.[133] This presumption depends on the requirement in GDPR Art 89(1) to ensure appropriate technical and organisational safeguards, such as pseudonymisation and access limitations. The EDPS has recently provided a Preliminary Opinion on data protection and scientific research ("**EDPS Preliminary Opinion on Scientific Research**")[134], whereby it takes an approach similar to the one applied by WP29 for presumption of compatibility under DPD Art 6(1)(b)[135]: "The presumption is not a general authorisation to further process data in all cases for historical, statistical or scientific purposes. Each case must be considered on its own merits and circumstances. But in principle personal data collected in the commercial or healthcare context, for example, may be further used for scientific research purposes, by the original or a new controller, if appropriate safeguards are in place."[136] Note that the EDPB intends to issue guidance on the 'horizontal and complex' conditions for the applicability of the 'presumption of compatibility' of further processing for archiving purposes in the public interest, scientific, historical research or statistical purposes, as provided for by the GDPR Article 5(1)(b).[137]

The presumption of compatibility has gained renewed attention in light of scientific research, particularly in the field of medical research, due to GDPR Rec 50, which reads as follows:

---

[130] DPD Art 6(1)(b).
[131] *Op cit.*, Article 29 Data Protection Working Party. Opinion 03/2013 on purpose limitation, 2013, p 13.
[132] *Ibid*.
[133] GDPR Art 5(1)(b).
[134] *Op. cit.*, European Data Protection Supervisor, A Preliminary Opinion on data protection and scientific research, 2020.
[135] *Op cit.*, Article 29 Data Protection Working Party. Opinion 03/2013 on purpose limitation, 2013, p 28.
[136] *Op. cit.*, European Data Protection Supervisor, A Preliminary Opinion on data protection and scientific research, 2020, p 22.
[137] *Ibid*, p 16.

> *The processing of personal data for purposes other than those for which the personal data were initially collected should be allowed only where the processing is compatible with the purposes for which the personal data were initially collected. In such a case, no legal basis separate from that which allowed the collection of the personal data is required. If the processing 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, Union or Member State law may determine and specify the tasks and purposes for which the further processing should be regarded as compatible and lawful. Further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes should be considered to be compatible lawful processing operations. The legal basis provided by Union or Member State law for the processing of personal data may also provide a legal basis for further processing. [emphasis added] […]*

The EDPS has highlighted a legal uncertainty concerning the need for a separate legal basis for secondary use of personal data for the purposes of scientific research.[138] This uncertainty has risen due to the phrase "where the processing is compatible with the purposes for which the personal data were initially collected […], no legal basis separate from that which allowed the collection of the personal data is required". According to EDPS, even though it "appears to assimilate purpose specification and lawfulness in the case of reuse for the purposes of scientific research", it is "not so much a blanket exemption to the separate steps set out in the Charter Article 8(2) - applicable to all circumstances - but rather advisory (hence 'should be considered to be compatible')."

At least one other interpretation of GDPR Rec 50 is possible – the aim of GDPR Rec 50 is simply to distinguish between "(initial) processing" and "further processing". The former – "(initial) processing" – signifies the activities that involve and directly follow after the initial collection of personal data, thus subjected to one and the same primary purpose which may be also covered by one and the same legal basis. The latter – "further processing" – refers to secondary use of the personal data initially collected for the primary purpose, in which case a secondary purpose is defined (for example, archiving in the public interest, scientific or historical research, and statistical purposes) and a new legal basis is required, even if the secondary purpose of processing is compatible with the primary purpose. According to this interpretation, further processing of personal data for whichever purposes requires a legal basis separate from that which allowed the collection of the personal data in the first place. For the sake of clarity, this is the interpretation that is adopted in the Sample DPIA. It concurs with the preliminary position expressed by the EDPB in its Preliminary Opinion on Scientific Research, in that the EDPB also does not see GDPR Rec 50 as exemption from the separate and cumulative requirements of purpose limitation and lawfulness. If, however, EDPB provides more guidance on the matter in the future, the interpretation adopted above may need to be revised.

[138] *Op. cit.*, European Data Protection Supervisor, A Preliminary Opinion on data protection and scientific research, 2020, p 23.

Another legal uncertainty revolving around GDPR Rec 50 concerns the need to carry out a compatibility test. Here, the EDPS has taken a conservative approach in the context of scientific research, in line with the earlier position of WP29 on the same matter[139], arguing that "in order to ensure respect for the rights of the data subject, the compatibility test under [GDPR Art 6(4)] should still be considered prior to the reuse of data for the purposes of scientific research, particularly where the data was originally collected for very different purposes or outside the area of scientific research."[140] Since the EDPS provided this interpretation in the context of scientific research, and not statistics, a question arises – is it obligatory under GDPR Art 6(4) to carry out a compatibility test prior to the reuse of mobile location data data for statistical purposes, considering that it was originally collected for different purposes outside the area of official statistics?

The EDPB has confirmed the relevance of the WP29 Opinion on Purpose Limitation guidance for the understanding of the principle of purpose limitation under DPD, although the WP29 Opinion is not adopted by the EDBP.[141] It thus makes sense to rely on the WP29 Opinion on Purpose Limitation when delineating the applicability of the presumption of compatibility between scientific research and statistical purposes. According to WP29, the presumption of compatibility in case of further processing for historical, statistical or scientific purposes contained in DPD Art 6(1)(b) "should not be read as providing an overall exception from the requirement of compatibility, and is not intended as a general authorization to further process data in all cases for historical, statistical or scientific purposes. Just like in any other case of further use, all relevant circumstances and factors must be taken into account when deciding what safeguards, if any, can be considered appropriate and sufficient. In addition, as in other situations, a separate test must be carried out to ensure that the processing has a legal basis in one of the grounds listed in Article 7 and complies with other relevant requirements of the Directive." [142]

Despite GDPR Rec 50 allowing ample room for interpretations, the WP29 clarification regarding the presumption of compatibility does not leave room for an interpretation whereby scientific research and statistical purposes could be treated differently in terms of compatibility of further processing. Both are subject to the requirement of compatibility and legal basis. However, thanks to the presumption of compatibility in both cases, a partial application of the full compatibility test may be justified, as long as appropriate safeguards are in place. As the EDPB intends to issue further guidance on the applicability of the presumption of compatibility of further processing for archiving purposes in the public interest, scientific, historical research or statistical purposes,[143] it would be redundant to speculate on its scope

[139] *Op cit.*, Article 29 Data Protection Working Party. Opinion 03/2013 on purpose limitation, 2013, p 28.
[140] *Op. cit.*, European Data Protection Supervisor, A Preliminary Opinion on data protection and scientific research, 2020, p 23.
[141] European Data Protection Board. Guidelines 4/2019 on Article 25 Data Protection by Design and by Default. Adopten on 20 October 2020, p 19. – Internet: https://edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_201904_dataprotection_by_design_and_by_default_v2.0_en.pdf (04.05.2021).
[142] *Op cit.*, Article 29 Data Protection Working Party. Opinion 03/2013 on purpose limitation, 2013, p 28.
[143] *Ibid*, p 16.

and contents any further. For the time being, it is assumed that an assessment of appropriate and sufficient safeguards will need to be carried out, as a minimum requirement of the presumption of compatibility. In addition, a separate analysis should be carried out to ensure a legal basis for the data processing.

iii. Lawfulness

The WP29 considered purpose specification and lawfulness to be two separate and cumulative requirements, based on its interpretation of Art 8 of the Charter.[144] This means that in addition to a specified and compatible purpose, the processing also has to have a legitimate basis.

Processing personal data for statistical purposes does not constitute a lawful basis for processing *per se*.[145] Such processing would have to rely on one of the legal conditions listed in GDPR Art 6.[146] Considering the nature of statistical purposes, GDPR Art 6 offers three main alternatives for processing personal data in this domain:

1) **consent** – the data subject has given consent to the processing of his or her personal data for one or more specific purposes[147];
2) **legal obligation** – processing is necessary for compliance with a legal obligation to which the controller is subject;[148]
3) **public interest/official authority** – processing 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.[149]

The other legal options provided in GDPR Art 6 have been ruled out for the following reasons:

4) **performance of a contract** – does not apply because official statistics does not concern offering goods or services to individuals;[150]
5) **vital interests** – does not apply because official statistics presumably does not prejudice the vital interests of a data subject or another natural person;[151]
6) **legitimate interests** – does not apply to processing carried out by public authorities (such as Eurostat and/or national statistical institutes) in the performance of their tasks.[152] The inapplicability of Art 6(1)(f) (legitimate

[144] See: Sec 7.1.4.ii above.
[145] Analogous to the conclusions of R. Ducato concerning processing personal data for scientific research purposes. Note that according to R. Ducato, Estonia is the only exception which has recognized research and official statistics as an autonomous legal basis. – R. Ducato. Data protection, scientific research, and the role of information. – Computer Law & Security Review, 37 (2020) 105412, p 7. – Internet: https://doi.org/10.1016/j.clsr.2020.105412 (04.04.2021).
[146] If the data being processed falls under one of the special categories of personal data, the controller shall also have to fulfill one of the requirements provided in GDPR Art 9(2).
[147] GDPR Art 6(1)(a).
[148] GDPR Art 6(1)(c).
[149] GDPR Art 6(1)(e).
[150] GDPR Art 6(1)(b).
[151] GDPR Art 6(1)(d).
[152] GDPR Art 6(1)(f), GDPR Art 6(1) last sentence.

interest) ground in case of processing for statistical purposes has also been confirmed by the EDPS.[153]

In other words, for a controller to process personal data for statistical purposes, it would have to either obtain the consent of all data subjects whose personal data it is processing (GDPR Art 6(1)(a)), be obliged to carry out such processing under the law (GDPR Art 6(1)(c)), or be vested with a public interest task or official authority under the law for which it is necessary to process personal data for statistical purposes (GDPR Art 6(1)(e)).

The EDPS has confirmed the above conclusion in the context of analysing data protection derogations for farm statistics, providing additional guidance on which legal options can be applied in which stages of the life cycle of official statistics:

1. **preparation of official statistics** – usually relies on consent, in cases where responses to survey questions are voluntary;[154]
2. **other cases** – processing of personal data is typically based on:
    a. legal obligation – requires the processing operations to follow from an explicit legal obligation;[155]
    b. public interest/official authority – in the absence of an explicit legal obligation, producing official statistics can be considered to be necessary for a performance of a task carried out in the public interest.[156]

The legal obligation and public interest task/official authority should be laid down as legal basis for processing by Union or Member State law. The Union or the Member State law shall meet an objective of public interest and be proportionate to the legitimate aim pursued. The purpose of the processing shall be determined in that legal basis or, as regards the processing for public interest/official authority, shall be necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller. That legal basis may contain specific provisions to adapt the application of rules of GDPR.[157]

Member States may also maintain or introduce more specific provisions to adapt the application of the rules of GDPR with regard to processing for compliance with the legal obligation and public interest task/official authority grounds by determining more precisely specific requirements for the processing and other measures to ensure lawful and fair processing including for other specific processing situations as

---

[153] *Op cit.*, European Data Protection Supervisor. EDPS Opinion on safeguards and derogations under Article 89 GDPR in the context of a proposal for a Regulation on integrated farm statistics, 2017, p 11.
[154] *Op cit.*, European Data Protection Supervisor, Opinion 10/2017 on integrated farm statistics, 2017, p 11.
[155] *Ibid.*, p 12.
[156] *Ibid*.
[157] *Inter alia:* the general conditions governing the lawfulness of processing by the controller; the types of data which are subject to the processing; the data subjects concerned; the entities to, and the purposes for which, the personal data may be disclosed; the purpose limitation; storage periods; and processing operations and processing procedures, including measures to ensure lawful and fair processing such as those for other specific processing situations as provided for in GDPR Chapter IX. – GDPR Art 6(3).

provided for in GDPR Chapter IX, including processing for statistical purposes in GDPR Art 89(1).[158] As a general requirement, EU or Member State law should, within the limits of GDPR, determine statistical content, control of access, specifications for the processing of personal data for statistical purposes and appropriate measures to safeguard the rights and freedoms of the data subject and for ensuring statistical confidentiality.[159]

[158] GDPR Art 6(2).
[159] GDPR Rec 162.

iv. Appropriate safeguards for processing for statistical purposes

Processing for statistical purposes enjoys a special regime within the GDPR[160], if appropriate safeguards for the rights and freedoms of data subjects are provided.[161] As concluded above, the implementation of appropriate safeguards is assumed to be the minimum requirement of applying the presumption of compatibility in case of processing for statistical purposes (see Section 7.1.4.ii.c) above).

The appropriate safeguards have to ensure that technical and organisational measures are in place in particular in order to ensure respect for the principle of data minimisation. They may include pseudonymisation but if the statistical purposes can be fulfilled by further processing which does not permit or no longer permits the identification of data subjects, those purposes shall be fulfilled in that manner.[162] The EDPS has also confirmed that "in cases where the purposes of the processing (or further processing) can be fulfilled without identification of the individuals, the controller must go beyond pseudonymisation and must ensure that individuals can no longer be identified."[163] More recently, the EDPS has emphasized in its Guidelines 04/2020 on the use of location data and contact tracing tools in the context of the COVID-19 outbreak[164] ("**EDPB Corona App Guidelines**") that preference should always be given to the processing of anonymised data, when it comes to using location data.[165] Hence, if there is a choice between pseudonymous vs deidentified or anonymous processing, the latter should be preferred whenever possible.

The basic idea of conditioning the application of the special regime on the implementation of appropriate safeguards is not new – it was present already in DPD Art 6(1)(b).[166] The further processing of data for historical, statistical and scientific research was allowed as long as the controller compensates for the change of purpose by implementing appropriate safeguards, in particular by ensuring that the data will not be used to support measures or decisions regarding any particular individuals.[167] Using the words of the WP29, "the purpose of the safeguards is typically to 'rule out' that the data will be used to support measures or decisions

---

[160] Analogous to the special regime for scientific research, which was created by adopting the GDPR. See: European Data Protection Supervisor. A Preliminary Opinion on data protection and scientific research. 6 January 2020, p 18 and footnote 102. – Internet: https://edps.europa.eu/data-protection/our-work/publications/opinions/preliminary-opinion-data-protection-and-scientific_en (08.04.2021).
[161] GDPR Art 89(1).
[162] *Ibid*.
[163] *Op cit.*, European Data Protection Supervisor. EDPS Opinion on safeguards and derogations under Article 89 GDPR in the context of a proposal for a Regulation on integrated farm statistics, 2017, p 13.
[164] European Data Protection Board. Guidelines 04/2020 on the use of location data and contact tracing tools in the context of the COVID-19 outbreak. Adopted on 21 April 2020. – Internet: https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-042020-use-location-data-and-contact-tracing_en (02.05.2021).
[165] *Ibid.*, p 5.
[166] *Op cit.*, Article 29 Data Protection Working Party. Opinion 03/2013 on purpose limitation, 2013, p 28.
[167] *Ibid*.

regarding any particular individual. The term 'rule out' suggests that the safeguards should indeed be strong enough to exclude or at least minimise the risks to the data subjects."[168] The WP29 further clarified that "any relevant impact on particular individuals – either negative or positive – should be avoided".[169]

WP29 Opinion on Purpose Limitation provides guidelines on how to conduct a compatibility assessment. According to this Opinion, appropriate safeguards are considered as a key factor to be considered as a last step of the compatibility assessment. They can serve as "compensation" for a change of purpose in case of further processing for historical, statistical or scientific purposes.[170] Among other things, this might require technical and/or organisational measures to ensure functional separation, such as partial or full anonymisation, pseudonymisation, and aggregation of data. When identifying appropriate safeguards to compensate for the change of purpose, the focus often lies with the notion of isolation.[171] WP29 has also considered privacy enhancing technologies as relevant to ensure that the data cannot be used to take decisions or other actions with respect to individuals (functional separation).[172]

According to the WP29 Opinion 05/2012 on Cloud Computing[173] ("**WP29 Opinion on Cloud Computing**"), isolation is a data protection goal which is meant to contribute to guaranteeing that data is not used beyond its initial purpose and to maintain confidentiality and integrity[174]. WP29 has specified in its Opinion on Cloud computing that achieving isolation requires:

1) adequate governance of the rights and roles for accessing personal data, which is reviewed on a regular basis (the implementation of roles with excessive privileges should be avoided, administrators and users must only be able to access information in accordance with the least privilege principle).
2) proper management of shared resources (if physical resources are shared between different customers).[175]

According to WP29 Opinion on Purpose Limitation, anonymisation is a key tool in achieving functional separation.[176] Along with pseudonymisation, aggregation, privacy enhancing technologies and other measures, anonymisation is particularly relevant in the context of further use for historical, statistical and scientific purposes.[177] However, since anonymisation does have its challenges and limits, once the first assessment has been completed in terms of the possibilities and limits

[168] *Ibid*. See also: DPD Rec 29.
[169] *Ibid*.
[170] *Op cit.*, Article 29 Data Protection Working Party. Opinion 03/2013 on purpose limitation, 2013, p 26.
[171] In Germany, the broader concept of "unlinkability", as promoted by the Confrerence of Data Protection Commissioners. – *Ibid.*, p 27, footnote 79.
[172] *Ibid.*, p 27.
[173] Article 29 Data Protection Working Paty. Opinion 05/2012 on Cloud Computing. Adopted July 1st 2012. 01037/12/EN WP196. – Internet: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2012/wp196_en.pdf (10.05.2021).
[174] *Ibid.*, p 15, sec 3.4.3.5 (Isolation (purpose limitation)).
[175] *Ibid.*, p 16, sec 3.4.3.5 (Isolation (purpose limitation)).
[176] *Op cit.*, Article 29 Data Protection Working Party. Opinion 03/2013 on purpose limitation, 2013, p 32.
[177] *Ibid.*, p 27.

of effective de-identification, the second step of applying additional safeguards will often need to follow.[178] WP29 has provided an example how anonymisation can be used as a tool for functional separation in case of further processing pseudonymous mobile location data:

> ***Example 15: Mobile phone locations help inform traffic calming measures***
>
> *The Department for Transport has asked a telecommunications company whether it can use the company's mobile phone location data. in order to calculate the speed at which the phones – and therefore the vehicles they are contained in – are moving over various routes. The mobile phone data reveals that speeding is common on certain stretches of road. This is then used to plan traffic-calming measures, which are later shown to have led to a significant reduction in road traffic accident fatalities in the area. The mobile phone data are effectively anonymised prior to disclosure to the Department of Transport to ensure that the risk of reidentification of the data subjects is minimal* [emphasis added]*. A careful impact assessment is made, penetration tests are carried out, and stakeholders are consulted. In this scenario we assume that all facts confirm very low or minimal risks of re-identification and relatively low impact on the data subjects if it nevertheless happens.*
>
> *This scenario requires a detailed compatibility assessment. Telecoms data initially collected for a specific purpose are now used for different (road traffic related) purposes. Most people would not commonly expect their data to be used in this way. This may give an initial strong indication that the purposes are incompatible. The relative sensitivity of the mobile location data collected may also support this assessment.*
>
> *However, in this case, prior to its use/disclosure for the secondary purpose, the data is effectively anonymised* [emphasis added]*. Therefore, although the two purposes are different, and provided the anonymisation is performed adequately (so the information no longer constitutes personal data or falls into a borderline zone with very low risks of re-identification) this reduces any concerns regarding incompatible processing. Nevertheless, additional safeguards, such as full transparency about the processing will be still recommended. In particular, if complete anonymisation cannot be ensured and some risks remain, this should be disclosed - as a rule, and unless an exemption under Article 13 could apply, informed consent will be required.*

v. The special regime for processing for statistical purposes

The EDPS has reiterated that "[t]he special regime [for scientific research] applies the usual principles such as lawfulness, purpose limitation and data subject rights, but permits some derogations from controller obligations. This includes the presumption of compatibility of processing for scientific research purposes of data collected in commercial and other contexts, provided appropriate safeguards are in

[178] *Ibid.*, p 32.

place."[179] According to R. Ducato[180], the structure of the special regime for scientific research consists of two branches:

1) exceptions to some data protection principles,
2) derogations to the exercising of a set of data subjects' rights
    a. laid down in the GDPR,
    b. can be introduced by EU or Member State law.

The EDPS statements and the structure of the special regime proposed by R. Ducato can be extended to the special regime for processing for statistical purposes by means of analogy because the GDPR norms they refer to apply both to scientific research and statistical purposes. Here is a more detailed overview of the elements of each branch adjusted to the context of processing for statistical purposes:

1. **exceptions to some data protection principles**:
    a. **purpose limitation** – further processing for statistical purposes shall not be considered to be incompatible with the initial purposes, provided that adequate safeguards are in place[181] (presumption of compatibility). Nevertheless, a compatibility assessment is still required, depending on the type of legal basis of further processing.[182]
    b. **storage limitation** – personal data may be stored for longer periods insofar as the personal data will be processed solely for statistical purposes subject to implementation of the appropriate technical and organisational measures in order to safeguard the rights and freedoms of the data subject;[183]
    c. **processing of special categories of personal data** – processing is necessary for statistical purposes based on Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject.[184]

2. **derogations to the exercising of a set of data subjects' rights**[185] – it is possible to distinguish between derogations that are[186]:
    a. laid down in the GDPR:
        i. **right to be informed (if the personal data have not been directly obtained from the data subject)** – the controller's obligation to provide information about the processing to the data subject, including in case of further processing of the personal data for other purposes, shall not apply if:

[179] European Data Protection Supervisor. A Preliminary Opinion on data protection and scientific research. 6 January 2020, p 2. – Internet: https://edps.europa.eu/data-protection/our-work/publications/opinions/preliminary-opinion-data-protection-and-scientific_en (14.04.2021).
[180] *Op. cit.*, R. Ducato. Data protection, scientific research, and the role of information. 2020, pp 5-6.
[181] GDPR Art 5(1)(b) and Art 89(1).
[182] GDPR Art 6(4).
[183] GDPR Art 5(1)(e) and Art 89(1).
[184] GDPR Art 9(2)(j) and Art 89(1).
[185] GDPR Art 89(2).
[186] *Op. cit.*, R. Ducato. Data protection, scientific research, and the role of information. 2020, pp 5-6.

1. the provision of such information proves impossible or
2. would involve a disproportionate effort, in particular for processing for statistical purposes, or
3. in so far as the obligation to provide information is likely to render impossible or seriously impair the achievement of the objectives of that processing.

In such cases the controller shall take appropriate measures to protect the data subject's rights and freedoms and legitimate interests, including making the information publicly available;[187]

ii. **right to erasure** – the controller's obligation to erase personal data without undue delay shall not apply to the extent that processing is necessary for statistical purposes in so far as the data subject's right to erasure is likely to render impossible or seriously impair the achievement of the objectives of that processing;[188]

iii. **right to object** – where personal data are processed for statistical purposes, the data subject, on grounds relating to his or her particular situation, shall not have the right to object to processing of personal data concerning him or her if the processing is necessary for the performance of a task carried out for reasons of public interest.[189]

b. **can be introduced by EU or Member State law**, in so far as such rights are likely to render impossible or seriously impair the achievement of the statistical purposes, and such derogations are necessary for the fulfilment of those purposes:[190]

i. right of access,[191]
ii. right to rectification,[192]
iii. right to restriction of processing,[193]
iv. right to object.[194]

Derogations introduced by EU or Member State law (p 2.b. above) must adhere to a "three-step-test" for statistical derogations – to verify whether there are legitimate grounds for the introduction of exceptions to data subjects' rights the following elements must be present cumulatively:

1) exercising the rights is likely to render impossible or seriously impair the achievement of statistical purposes;
2) the derogations must be necessary for the fulfilment of statistical purposes;
3) appropriate safeguards for the rights and freedoms of the data subject must be adopted.[195]

[187] GDPR Art 14(5)(b) and Art 89(1).
[188] GDPR Art 17(3)(d) and Art 89(1).
[189] GDPR Art 21(6) and Art 89(1).
[190] GDPR Art 89(2).
[191] GDPR Art 15 and Art 89(2).
[192] GDPR Art 16 and Art 89(2).
[193] GDPR Art 18 and Art 89(2).
[194] GDPR Art 21 and Art 89(2).
[195] Analogous to the conclusions of R. Ducato concerning processing personal data for scientific research purposes. - *Op. cit.*, R. Ducato. Data protection, scientific research, and the role of information. 2020, p 7.

7.2. Statistics law

7.2.1. Overview

As outline above, the EU data protection law distinguishes official statistics as a subset of statistics in general and is further divided into official European statistics and official national statistics. According to national and EU laws on official statistics, citizens and businesses are usually obliged to disclose data to the relevant statistics authorities. "Officials working in statistics bureaus are bound by special professional secrecy obligations which must be complied with properly, as they are essential for the high-level of citizen trust necessary if data are to be made available to the statistics authorities."[196]

The legal framework for official European statistics derives from the TFEU. According to the TFEU, the European Parliament and the Council have the right to adopt measures for the production of statistics where necessary for the performance of the activities of the Union.[197] TFEU also lays out the main principles to which the production of Union statistics must conform to: impartiality, reliability, objectivity, scientific independence, cost-effectiveness and statistical confidentiality; it shall not entail excessive burdens on economic operators.[198]

Currently, the EU and its Member States have shared competence in the field of statistics.[199] To this end, the RES provides for a cooperation mechanism between the EU and Member State level – it establishes the ESS as the partnership between the Eurostat, NSIs and other national authorities responsible in each Member State for the development, production and dissemination of European statistics.[200] The basic principles and rules for how the ESS should function, have been set out in the RES. However, it does not specify which statistics should be produced – this is a matter for sector-specific legislation.[201] European statistics are usually based on national data produced and disseminated by the national statistical authorities of all Member States.[202] In practice, Eurostat relies mostly on the NSIs regarding the production of statistics and assuring its quality.[203]

---

[196] GDPR Art 90; Handbook on European data protection law. 2018 edition. Luxembourg: Publications Office of the European Union, 2018, p 340. – Internet: https://www.echr.coe.int/Documents/Handbook_data_protection_ENG.pdf (12.04.2021).
[197] TFEU Art 338(1).
[198] TFEU Art 338(2).
[199] A. V. Georgiou. A New Statistical System for the European Union. Bruegel Essay and Lecture Series. Bruegel, 2018, pp 141-142. – In the Internet: https://www.bruegel.org/wp-content/uploads/2018/12/A-NEW-STATISTICAL-SYSTEM-FINAL.pdf (04.12.2020).
[200] RES Art 4.
[201] Eurostat. European Commission. Legal framework for European statistics. The Statistical Law. 2010 edition. Luxembourg: Publications Office of the European Union, 2010. – In the Internet: https://ec.europa.eu/eurostat/web/products-statistical-books/-/KS-31-09-254 (31.01.2021).
[202] RES Rec (15).
[203] There is an ongoing academic and policy discussion regarding the need to further integrate the ESS in order to assure the quality of European statistics. – A. V. Georgiou. A New Statistical System for the European Union. Bruegel Essay and Lecture Series. Bruegel, 2018. – In the Internet: https://www.bruegel.org/wp-content/uploads/2018/12/A-NEW-STATISTICAL-SYSTEM-FINAL.pdf (04.12.2020); A. Schout, A. Mijs. The governance of the ESS. Coordinating expectations. Clingendael Report, December 2016.

The legal framework for the development, production and dissemination of European statistics is based on RES.[204] RES defines European statistics as "relevant statistics necessary for the performance of the activities of the Community", which are determined in the European statistical programme and shall be developed, produced and disseminated in conformity with the statistical principles as set out in the TFEU and further elaborated in the European Statistics Code of Practice[205] ("**ESCoP**").[206] European statistics are developed, produced and disseminated by both the ESS and the ESCB but under separate legal frameworks reflecting their respective governance structures.[207]

The legal framework of European statistics is complemented by a self-regulatory common quality framework of the ESS, which consists of the ESCoP[208], Quality Assurance Framework of the European Statistical System[209] ("**QAF ESS**") and the general quality management principles (such as continuous interaction with users, commitment of leadership, partnership, staff satisfaction, continuous improvement, integration and harmonisation).[210] ESCoP shall give ethical guidance on how to perform official statistics[211] – it sets the standard for developing, producing and disseminating European statistics, along the lines of the institutional environment, statistical processes and statistical output.[212] QAF ESS breaks further down the ESCoP – it identifies possible methods, tools and good practices that can provide guidance and evidence for the implementation of the ESCoP.[213] Statistical authorities, comprising the Eurostat, the NSIs and other national authorities responsible for the development, production and dissemination of European statistics are self-committed to continuously develop, produce and disseminate high-quality European statistics and services in order to sustainably provide value to its users, as demonstrated by the Quality Declaration of the European Statistical System[214] ("**QD ESS**").[215]

– In the Internet: https://www.clingendael.org/sites/default/files/pdfs/The_Governance_of_the_ESS.pdf (04.12.2020); On a related topic: W. J. Radermacher. Official Statistics 4.0. Verified Facts for People in the 21st Century. Springer, Cham, 2020. – In the Internet: https://link.springer.com/book/10.1007/978-3-030-31492-7#about (04.12.2020).

[204] RES Art 1(1).

[205] European Statistics Code of Practice. For the National Statistical Authorities and Eurostat (EU statistical authority). Adopted by the European Statistical System Committee, 16th November 2017 – Internet: https://ec.europa.eu/eurostat/web/products-catalogues/-/ks-02-18-142 (12.04.2021).

[206] RES Art 1(2).

[207] RES Rec (8).

[208] Based on RES Art 11.

[209] Quality Assurance Framework of the European Statistical System. Version 2.0. – Internet: https://ec.europa.eu/eurostat/documents/64157/4392716/ESS-QAF-V2.0-final.pdf (12.04.2021).

[210] European Statistics Code of Practice. For the National Statistical Authorities and Eurostat (EU statistical authority). Adopted by the European Statistical System Committee. 16th November 2017, p 5 – Internet: https://ec.europa.eu/eurostat/web/products-catalogues/-/ks-02-18-142 (12.04.2021).

[211] *Op cit.,* Handbook on European data protection law, 2018, p 341, footnote 958.

[212] Eurostat web page. Quality. Overview. Internet: https://ec.europa.eu/eurostat/web/quality (12.04.2021).

[213] *Ibid*.

[214] Quality Declaration of the European Statistical System. September 2016. – Internet: https://ec.europa.eu/eurostat/web/products-catalogues/-/KS-02-17-428 (12.04.2021).

[215] *Op cit.,* European Statistics Code of Practice, 2017, p 5.

The QD ESS also envisions a forward-looking approach, whereby the NSIs and Eurostat not only improve the quality of their products and services by continuously modernising, innovating and compiling new indicators, but also attempt to anticipate emerging phenomena and needs with their users.[216]

7.2.2. Legal definition of statistical purposes in the context of European statistics

Compared to GDPR, RES sets out its own terminology relating to European statistics. According to RES Art 3(1)(1), "statistics" means quantitative and qualitative, aggregated and representative information characterising a collective phenomenon in a considered population. RES Art 3(1)(8) provides the meaning of "use for statistical purposes" as the exclusive use for the development and production of statistical results and analyses. Here, "development" means the activities aiming at setting up, strengthening and improving the statistical methods, standards and procedures used for the production and dissemination of statistics as well as at designing new statistics and indicators;[217] "production" means all the activities related to the collection, storage, processing, and analysis necessary for compiling statistics.

Due to its guiding function, the ESCoP and its Glossary can also act as a source for interpreting the definitions and rules provided in the RES. According to the Glossary, the key terms used in the ESCoP are defined as follows:

- "**European statistics**" means relevant statistics that are necessary for the performance of the activities of the EU and are defined in the relevant statistical programme.[218]
- "**official statistics**" means statistics describing on a representative basis phenomena of public interest to policy makers, the economic agents and the public at large. They are developed, produced and disseminated by the statistical authorities in compliance with the provisions of the Union and national law and the ESCoP / National Codes of Practice. They shall be referred to as 'official statistics' in the statistical programme.[219]
- "**statistical authorities**" are defined as the bodies responsible for the development, production and dissemination of European statistics. They: a) exercise public authority based on national law; b) have production of statistics included among their tasks in the respective basic act; c) have clearly been given the responsibility at the national level for the production of a specific and identifiable part of European statistics.[220]

7.2.3. Data protection and statistical confidentiality

[216] Quality Declaration of the European Statistical System. September 2016, p 2 – Internet: https://ec.europa.eu/eurostat/web/products-catalogues/-/KS-02-17-428 (12.04.2021).
[217] RES Art 3 p 2.
[218] Glossary. Defining the main terms used in the European Statistics Code of Practice, as adopted by the ESSC of November 2017, p 1 – Internet: https://ec.europa.eu/eurostat/documents/4031688/9439112/Glossary/ (12.04.2021).
[219] *Ibid*.
[220] *Ibid*.

Statistical confidentiality is one of the main statistical principles outlined in the TFEU.[221] RES provides further specifications on statistical confidentiality for European statistics.[222] It defines the meaning of the principle of statistical confidentiality as the "protection of confidential data related to single statistical units which are obtained directly for statistical purposes or indirectly from administrative or other sources and implying the prohibition of use for non-statistical purposes of the data obtained and of their unlawful disclosure".[223] In other words, statistical confidentiality requires that confidential data are exclusively used for statistical purposes and their unlawful disclosure is prevented.[224]

According to RES, "confidential data" means data which allow statistical units to be identified, either directly or indirectly, thereby disclosing individual information. To determine whether a statistical unit is identifiable, account shall be taken of all relevant means that might reasonably be used by a third party to identify the statistical unit.[225] The EDPS has analysed the parallelism between the concept of confidential data and the one of personal data. The conclusion was that statistical confidentiality and data protection, although presenting similarities in wordings, cover two different concepts. Due to the possibility of confusion between the two notions, the EDPS underlined the need to clearly assess the differences between data protection and statistical confidentiality.[226] It clarified that the notion "personal data" relates exclusively to natural persons, whereas the definition of "statistical confidentiality" relates also to households, economic operators and other undertakings, in addition to natural persons.[227]

Along the same lines, the notion of anonymity also has a different scope in data protection law and in statistics law. From a data protection view, the notion of anonymity would cover data that are no longer identifiable (previously DPD Rec 26, now GDPR Rec 26).[228] From a statistical point of view, anonymous data are data for which no direct identification is possible, implying that indirect identification of data would still qualify these data as anonymous.[229] Therefore, the threshold for treating personal data as anonymous is different in each case – it is lower under statistics law (includes indirectly identifiable data) and higher under data protection law (does not include identifiable data). The EDPS has emphasized that "in order to avoid possible misunderstandings when using these notions, the context and legal

[221] TFEU Art 338(2).
[222] GDPR Rec 163.
[223] RES Art 2(1)(e).
[224] RES Art 20(1).
[225] RES Art 3(7).
[226] EDPS. Opinion of the European Data Protection Supervisor on the proposal for a Regulation of the European Parliament and of the Council on European Statistics (COM(2007) 625 final), 2008/C 308/01, 03.12.2008, Sec 17, p 3. – Internet: https://edps.europa.eu/data-protection/our-work/publications/opinions/european-statistics_en (14.04.2021).
[227] *Ibid.*, Sec 19, p 3.
[228] EDPS. Opinion of the European Data Protection Supervisor on the proposal for a Regulation of the European Parliament and of the Council on Community statistics on public health and health and safety at work (COM(2007) 46 final), 2007/C 295/01, Brussels, 5 September 2007, Sec 19 – Internet: https://edps.europa.eu/data-protection/our-work/publications/opinions/community-statistics-health-data_en (25.08.2021).
[229] *Ibid*.

framework in which these notions are being used should be always clearly and precisely defined."[230]

In the context of statistical confidentiality, the concept of "functional separation" is of particular relevance. It means that data used for statistical purposes or other research purposes should not be available to support measures or decisions that are taken with regard to the individual data subjects concerned (unless specifically authorized by the individuals concerned).[231] "Functional separation and statistical confidentiality require organisations to put in place technical and organizational measures to ensure that personal data processed for statistical purposes cannot be used for non-statistical purposes."[232] This requirement includes the prohibition of using confidential data used for statistics for informing decisions or measures that would directly affect the individuals concerned (e.g. the responses of an individual to a statistical survey cannot be used by tax authorities to determine the respondent's tax liability).[233]

The rules and measures to ensure the principle of statistical confidentiality are provided in RES Chapter V "Statistical Confidentiality". In the context of functional separation, a distinction is made between two types of sources for confidential data related to single statistical units:

1) **data obtained directly for statistical purposes** – this data shall be used by the NSIs and other national authorities and by Eurostat exclusively for statistical purposes[234] and, hence, may not be used for any other purpose.
2) **data obtained indirectly from administrative or other sources** – this data was initially collected for non-statistical purposes, but it shall be available for further statistical use.[235]

7.2.4. Data protection and statistical quality

According to RES Rec 25, the availability of confidential data for the needs of the ESS is of particular importance in order to maximise the benefits of the data with the aim of increasing the quality of European statistics and to ensure a flexible response to the newly emerging Community statistical needs.[236] Therefore, the goal of making confidential data available for NSIs and other members of the ESS is to ensure statistical quality and respond to new statistical needs.

7.3. Electronic communications law

[230] *Ibid.*
[231] *Op cit*, Article 29 Data Protection Working Party, Opinion 03/2013 on purpose limitation, 2013, p 30.
[232] *Op cit*., European Data Protection Supervisor. EDPS Opinion on safeguards and derogations under Article 89 GDPR in the context of a proposal for a Regulation on integrated farm statistics, 2017, p 13.
[233] *Ibid*.
[234] Unless the statistical unit has unambiguously given its consent to the use for any other purposes. – RES Art 20(2).
[235] *Op cit.,* Handbook on European data protection law, 2018, p 341.
[236] RES Rec 25.

7.3.1. Overview

The telecom sector, including electronic communications, is one of the most regulated industries in the EU. Electronic communications services generate traffic data and location data, which may involve personal data processing, insofar as they relate to natural persons. Therefore, the telecom sector regulations provide some specific obligations concerning the protection of fundamental rights and freedoms *vis-à-vis* the processing of personal data regarding electronic communications. Such specific rules are currently laid out in the ePD, which may be replaced by a new and directly applicable regulation in the near future.[237]

The ePD Art 5(1) lays out the obligation to ensure confidentiality of communications and the related traffic data:

> *(1) Member States shall ensure the confidentiality of communications and the related traffic data by means of a public communications network and publicly available electronic communications services, through national legislation. In particular, they shall prohibit listening, tapping, storage or other kinds of interception or surveillance of communications and the related traffic data by persons other than users, without the consent of the users concerned, except when legally authorised to do so in accordance with Article 15(1). This paragraph shall not prevent technical storage which is necessary for the conveyance of a communication without prejudice to the principle of confidentiality.*[238]

The customer relationship between MNOs (electronic communications service providers) and the Subscribers triggers the material scope of both the ePD and the GDPR.[239] If the ePD renders more specific rules than the GDPR, then the specific provisions of the ePD shall, as *lex specialis*, take precedence over the more general provisions of the GDPR. Any processing of personal data which is not specifically governed by the ePD (or for which the ePD does not contain a "special rule"), remains subject to the provisions of the GDPR.[240]

In general, the processing of personal data can be justified on the basis of each of the lawful grounds mentioned in GDPR Art 6. However, the full range of possible lawful grounds provided by GDPR Art 6 cannot be applied by an MNO to processing of traffic or location data, because ePD explicitly limits the conditions in which such data, including personal data, may be processed.[241] The specific situations where the ePD particularises the provisions of the GDPR are as follows:

---

[237] European Commission. Shaping Europe's digital future. Proposal for an ePrivacy Regulation – Internet: https://digital-strategy.ec.europa.eu/en/policies/eprivacy-regulation (18.05.2021).
[238] ePD Art 5(1).
[239] European Data Protection Board. Opinion 5/2019 on the interplay between the ePrivacy Directive and the GDPR, in particular regarding the competence, tasks and powers of data protection authorities. Adopted on 12 March 2019, Sec 34, p 12. – In the Internet: https://edpb.europa.eu/our-work-tools/our-documents/styrelsens-yttrande-art-64/opinion-52019-interplay-between-eprivacy_en (20.01.2021).
[240] *Ibid.*, Sec 38, p 13.
[241] *Ibid.*, Sec 39, p 13.

1) **traffic data**: ePD Art 6 concerning the processing of so-called traffic data limits the conditions in which traffic data, including personal data, may be processed. Here, the more specific provisions of the ePD Art 6 must take precedence over the more general provisions of the GDPR, but it does not curtail the applications of other provisions of the GDPR, such as the rights of the data subject or the requirement that processing of personal data must be lawful and fair.[242] In other words, if a type of data processing is not allowed under ePD Art 6, there cannot be a legal ground for it in GDPR Art 6.[243]

2) **terminal equipment information**: ePD Art 5(3) provides that, as a rule, prior consent is required for the storing of information, or the gaining of access to information already stored, in the terminal equipment of a subscriber or user. If such information constitutes personal data, then ePD Art 5(3) shall take precedence over GDPR Art 6 with regard to the activity of storing or gaining access to this information. This means that where ePD Art 5(3) requires consent for the specific actions it describes, the controller cannot rely on the full range of possible lawful grounds provided by GDPR Art 6.[244]

3) **location data and electronic contact details**: The outcome is similar in the interplay between ePD Art-s 9 (location data other than traffic data) and 13 (unsolicited communications), on the one hand, and GDPR Art 6, on the other hand. For example, according to ePD Art 9(1), location data other than traffic data may only be processed when (a) they are made anonymous, or (b) with the consent of the users or subscribers to the extent and for the duration necessary for the provision of a value added service. Similarly to the case of traffic data and terminal equipment information above, this means that where ePD Art 9(1) requires anonymisation of location data other than traffic data or a consent for the provision of value added services, the controller cannot rely on the full range of possible lawful grounds provided by GDPR Art 6.[245] In conclusion, if a type of data processing is not allowed under ePD Art 9 or 13, there cannot be a legal ground for it in GDPR Art 6.

7.3.2. Conditions for further processing of mobile location data

Due to the *lex specialis* nature of the specific provisions of the ePD over the more general provisions of the GDPR explained above, there is a limited choice of options for further processing of mobile location data, which the MNOs collected for their legitimate purposes recognised under the ePD. Further, the legal basis for such

---

[242] *Ibid*.
[243] Note that the WP29 came to the same conclusion in its analysis of the interplay between ePD Art 6 and DPD (GDPR predecessor) Art 7 in its Opinion on Anonymisation Techniques (p 8). However, the EDPB has not officially endorsed the WP29 Opinion on Anonymisation Techniques up until now. - European Data Protection Board. GDPR: Guidelines, Recommendations, Best Practices. Endorsement of GDPR WP29 Documents. – In the Internet: https://edpb.europa.eu/our-work-tools/general-guidance/gdpr-guidelines-recommendations-best-practices_en (21.01.2021).
[244] O*p. cit.*, European Data Protection Board, Opinion 5/2019 on the interplay between the ePrivacy Directive and the GDPR, 2019, Sec 40, pp 13-14.
[245] O*p. cit.*, European Data Protection Board, Opinion 5/2019 on the interplay between the ePrivacy Directive and the GDPR, 2019, Sec 40, p 14.

further processing would have to be established in the relevant EU or Member State law[246], presuming the relevant norms of ePD do not have direct effect.[247]

Taking into account that the Sample Use Case deals with mobile location data, the most relevant option for allowing its further processing is currently provided under ePD Art 9, which reads as follows:

> *1. Where location data other than traffic data, relating to users or subscribers of public communications networks or publicly available electronic communications services, can be processed, such data may only be processed when they are made anonymous, or with the consent of the users or subscribers to the extent and for the duration necessary for the provision of a value added service* [emphasis added]. *The service provider must inform the users or subscribers, prior to obtaining their consent, of the type of location data other than traffic data which will be processed, of the purposes and duration of the processing and whether the data will be transmitted to a third party for the purpose of providing the value added service. Users or subscribers shall be given the possibility to withdraw their consent for the processing of location data other than traffic data at any time.*
>
> *2. Where consent of the users or subscribers has been obtained for the processing of location data other than traffic data, the user or subscriber must continue to have the possibility, using a simple means and free of charge, of temporarily refusing the processing of such data for each connection to the network or for each transmission of a communication.*
>
> *3. Processing of location data other than traffic data in accordance with paragraphs 1 and 2 must be restricted to persons acting under the authority of the provider of the public communications network or publicly available communications service* [emphasis added] *or of the third party providing the value added service, and must be restricted to what is necessary for the purposes of providing the value added service.*

Based on the wording of ePD Art 9(1) and Art 9(3), there seem to be two alternatives for further processing of mobile location data:

---

[246] For example, the EDPS has recently declared, in the context of analysing the legitimacy of new legal measures in the field of statistics relating to persons and households, that „[f]urther legislative measures in the field of national or Union law governing statistics, however, will likely to be still required, in order to allow more widespread use of big data in statistics in a way compatible with applicable data protection law. /--/ The current Proposal should not provide the illusion that Article 8 itself is providing a sufficient legal basis for using big data for the purposes of the Proposal. It is essential that the recitals and Article 8, combined, make it clear that any such use of big data sources is subject to applicable data protection law, including the need for an appropriate legal basis under Article 6 of the GDPR." – European Data Protection Supervisor. Opinion 2/2017. EDPS Opinion on the proposed common framework for European statistics relating to persons and households. 1 March 2017, sec 19-20. – Internet: https://edps.europa.eu/data-protection/our-work/publications/opinions/european-statistics-0_en (13.05.2021).

[247] The issue of direct applicability of a legal norm in the ePD needs further analysis. It is not within the scope of this document.

1) **for unlimited purposes**, as long as the data is made anonymous and the making anonymous is carried out by persons under the authority of the MNO;
2) **for the purposes of providing value added services**, as long as the extent and duration of the processing is necessary for such purposes, Subscribers have consented to it beforehand and the processing is carried out by persons under the authority of the MNO or the third party providing the value added services.

7.3.3. Timing of the "making anonymous" step

The "making anonymous" requirement has been analysed by the ECJ: "As regards location data other than traffic data, Article 9(1) of [ePD] provides that that data may be processed only subject to certain conditions and after [emphasis added] it has been made anonymous or the consent of the users or subscribers obtained."[248] What is curious about the ECJ position is that it seems to have attached a temporal aspect to the ePD Art 9(1) first alternative, requiring that the "making anonymous" step be carried out before further data processing. The temporal aspect is of crucial importance in the context of the Sample DPIA, because if the ECJ position is upheld, then the most feasible legal route to processing pseudonymous mobile location data by means of the Solution for producing official statistics is to carry out the processing in two steps:

1. **making the data anonymous** – the mobile location data has to be made anonymous before its substantive analysis,
2. **statistical analysis** – once the requirement of making the mobile location data anonymous is fulfilled, the substantive analysis of the mobile location data can be carried out.

For these reasons, the temporal aspect of the "make anonymous" requirement of ePD Art 9(1) first alternative needs to be clarified – is it acceptable under ePD Art 9(1) first alternative to make data anonymous along with statistical analysis, i.e. carry out both types of processing at the same time?

The text of ePD Art 9(1) first alternative does not indicate any clear conditions in terms of timing:

> *(1) Where location data other than traffic data, relating to users or subscribers of public communications networks or publicly available electronic communications services, can be processed, such data may only be processed when they are made anonymous* [emphasis added] *[…]"*[249]

Even if a chronological order of activities could be established, whereby the "making the data anonymous" step would have to precede the "statistical analysis" step, it is crucial to understand that the act of "making anonymous" is a kind of processing by itself. All the current well-known techniques of anonymisation – data randomization, aggregation, suppression and generalization – require some analysis of the

[248] European Court of Justice. Judgment of the Court (Grand Chamber) of 21 December 2016. Tele2 Sverige AB v Post- och telestyrelsen and Secretary of State for the Home Department v Tom Watson and Others. Joined Cases C-203/15 and C-698/15, sec 86. – Internet: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=ecli%3AECLI%3AEU%3AC%3A2016%3A970 (13.05.2021).
[249] ePD

underlying personal data in order to effectively eliminate linkages between the data and the relevant individuals. At the same time, all these techniques are aimed at preserving the usefulness of the anonymous data – the specific method and implementation of anonymisation depends on the context and purposes of using the resulting anonymous data, i.e. what kind of data is needed to conduct the further processing. Therefore, making data anonymous is a context-specific processing – it cannot always be carried out in two clearly distinguishable steps, where "making anonymous" happens before any further processing.

In light of the above, the temporal aspect to the ePD Art 9(1) first alternative requires interpretation. The ECJ has clarified the purpose and scope of ePD Art 9(1) as follows:

> 87 *The scope of Article 5, Article 6 and Article 9(1) of Directive 2002/58, which seek to ensure the confidentiality of communications and related data, and to minimise the risks of misuse* [emphasis added]*, must moreover be assessed in the light of recital 30 of that directive, which states: 'Systems for the provision of electronic communications networks and services should be designed to limit the amount of personal data necessary to a strict minimum* [emphasis added]*'.*[250]

Based on this, the primary goal of ePD Art 9(1) first alternative is to ensure confidentiality of communications and related data, and to minimise the risks of misuse. The best measure to achieve this goal – absolute minimisation of risks – is simply to delete the relevant data, instead of "making it anonymous". However, as the ECJ has pointed out, the scope of ePD Art 9(1) has to be assessed in light of ePD Rec 30, which states as follows:

> *(30) Systems for the provision of electronic communications networks and services should be designed to limit the amount of personal data necessary to a strict minimum* [emphasis added]*. Any activities related to the provision of the electronic communications service that go beyond the transmission of a communication and the billing thereof should be based on aggregated, traffic data that cannot be related to subscribers or users. Where such activities cannot be based on aggregated data, they should be considered as value added services for which the consent of the subscriber is required.*[251]

Therefore, ePD Rec 30 delineates the scope of application of ePD Art 9(1) – it adds the dimension of necessity. For the purposes of the Sample DPIA, this can be interpreted so that "making anonymous" and any further processing (e.g. statistical analysis) should be separate, consecutive processing steps only if this is required "to limit the amount of personal data necessary to a strict minimum." If, however, the underlying personal data loses its value and utility as a result of such separation, then it may necessitate to carry out the "making anonymous" and any further processing (e.g. statistical analysis) steps simultaneously. Throughout the following analysis, the authors take the position that this interpretation is not prohibited, *per se*.

---

[250] *Op. cit.*, European Court of Justice. Judgment of the Court (Grand Chamber) of 21 December 2016. Tele2 Sverige AB v Post- och telestyrelsen and Secretary of State for the Home Department v Tom Watson and Others. Joined Cases C-203/15 and C-698/15, sec 87.
[251] ePD Rec 30.

Furthermore, ePD Rec 9 also emphasizes the objectives of minimising the processing of personal data and of using anonymous or pseudonymous data where possible:

> *(9) The Member States, providers and users concerned, together with the competent Community bodies, should cooperate in introducing and developing the relevant technologies where this is necessary to apply the guarantees provided for by this Directive and taking particular account of the objectives of minimising the processing of personal data and of using anonymous or pseudonymous data where possible* [emphasis added].[252]

This indicates that "using anonymous data" is dependent on its feasibility – if "using anonymous data" would not make sense for the purposes of the further processing, it is reasonable to presume that "making data anonymous" and further processing may be conducted in parallel.

The WP29 has also confirmed that the process of anonymisation is context-dependent and different measures may be applied to ensure anonymity, for example:

> *[…] the assessment whether the data allow identification of an individual, and whether the information can be considered as anonymous or not depends on the circumstances, and a case-by-case analysis should be carried out with particular reference to the extent that the means are likely reasonably to be used for identification as described in Recital 26.*[253]
>
> *[…] removing directly identifying elements in itself is not enough to ensure that identification of the data subject is no longer possible. It will often be necessary to take additional measures to prevent identification, […] depending on the context and purposes of the processing for which the anonymised data are intended.*[254]

For practical reasons, and for the purposes of the Sample DPIA, the authors of this analysis presume that there is no requirement to complete the "making anonymous" step before further processing according to ePD Art 9(1) first alternative. Rather, the need to conduct the "making anonymous" step before further processing is dependent on the context and purposes of the processing for which the anonymised data are intended. Otherwise, the lawmaker and data protection regulators would have to provide clear guidelines for distinguishing the "making anonymous" step from other types of further processing and understanding when the "making anonymous" step is complete to allow further processing to continue. Due to the context-specific nature of "making anonymous", such guidelines would have little practical value as

---

[252] ePD Rec 9.
[253] *Op. cit.*, Article 29 Data Protection Working Party. Opinion 4/2007 on the concept of personal data, 2007, p 21.
[254] Article 29 Data Protection Working Party. Opinion 05/2014 on Anonymisation Techniques. Adopted on 10 April 2014. 0829/14/EN WP216. – Internet: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf (30.04.2021).

there is a grey area between “making anonymous” and other types of further processing.

7.3.4. Statistical analysis as a value added service

For the purposes of this DPIA, it deserves some attention to clarify whether the ePD Art 9(1) second alternative – processing for the purposes of providing value added services – may cover processing for statistical purposes, including official statistics. This presumes that statistical analysis is considered as a value added service.

According to ePD Art 2(g), ““value added service” means any service which requires the processing of traffic data or location data other than traffic data beyond what is necessary for the transmission of a communication or the billing thereof” – the legal definition is broad enough to incorporate any type of processing with some valuable results. However, when analysing the examples of value added services in the Recitals of ePD, it becomes clear that the legislator had a much narrower understanding in mind – value added services may, for example, consist of advice on least expensive tariff packages, route guidance, traffic information, weather forecasts and tourist information[255] or services providing individualised traffic information and guidance to drivers.[256] This means that value added services are meant to create direct value to the Subscribers, who have consented to their data being processed for such purposes.

Such direct value cannot be equalled to the public benefits of official statistics. In addition, the ECJ has confirmed, at least in the context of traffic data, that the ePD provisions concerning, *inter alia*, value-added services, “do not concern the communication of that data to persons other than those acting under the authority of the [MNO]”[257]. In other words, value added services can be provided only under the authorisation of the MNO, which is not the case for producing official statistics.

The second alternative also includes a requirement that the processing for the purposes of providing value added services must be based on the Subscriber’s consent as a legal basis for personal data processing. Essentially, ePD Art 9(1) second alternative has determined data subject consent under GDPR Art 6(1) as the only suitable legal basis for processing location data for the purposes of providing value added services.

Due to the above considerations, this second alternative for further processing of mobile location data for providing value added services is not applicable for the purposes of official statistics.

[255] ePD Rec 18.
[256] ePD Rec 35.
[257] European Court of Justice. Judgment of the Court (Grand Chamber) of 29 January 2008. Productores de Música de España (Promusicae) v Telefónica de España SAU. Case C-275/06, European Court Reports 2008 I-00271, sec 48. – Internet: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=ecli%3AECLI%3AEU%3AC%3A2008%3A54 (13.05.2021).

7.3.5. Interim conclusion

The only way that mobile location data can be further processed according to EU law *de lege lata*, other than for providing value-added services to the Subscribers, is provided in ePD Art 9(1) first alternative. The conditions for applying the first alternative are as follows:

1. the mobile location data must be made anonymous. Presumably, there is no requirement to complete the "making anonymous" step before further processing according to ePD Art 9(1) first alternative.
2. the legal basis for making mobile location data anonymous is unspecified, i.e. here the ePD has not limited the choice of legal bases provided in GDPR Art 6;
3. the making anonymous must be carried out by persons under the authority of the MNO.

8. Legal analysis

8.1. Documented tasks and issues

According to the DPIA Methodology, the last step in the preparation stage of a DPIA after presenting the applicable laws is to document the results of the preparation stage, following a standardized procedure in the form of a scoping report.[258] *For the purposes of the Project, the present document as a whole fulfils the function of a scoping report. However, in addition, the present document includes a legal analysis chapter in order to analyse and solve the legal conflicts and uncertainty identified in the course of mapping the legal requirements. Without addressing these issues, they may cause further delays and hurdles before the Solution can be adopted in practice. The conclusions of the legal analysis will be used as input for the Evaluation Report.*

8.1.1. The two-sided nature of the Sample Use Case

Processing of mobile location data for the purposes of official statistics can be looked at from two perspectives:

1) on the one side, there is the NSI who wants access to new types of quality input data in order to live up to its public task as a valuable knowledge provider for the society and, potentially, extend its service portfolio in the data analytics market;
2) on the other side, there is the MNO who wants to add value to its existing mobile location data while not jeopardising the trust of its Subscribers.

In a typical statistical analysis process such as that employed in the Sample Use Case, the NSI needs to keep the input data re-linkable during the preparation phase in order to combine and enrich statistical data by linking various datasets. At a later phase, once re-linking is no longer necessary for the statistical purposes sought, the keys are usually destroyed and additional means are used to ensure statistical confidentiality. We shall refer to this as the **"first analyse, then make anonymous"** approach, required by statistics laws to be applied by the NSI due to the principle of statistical confidentiality.

According to ePD Art 9(1)[259], mobile location data can be further processed for any and all purposes *de lege lata* only if it is "made anonymous". This can be in direct conflict with the "first analyse, then make anonymous" approach, as long as the "make anonymous" and "analyse" functions are seen as necessarily distinct steps where the "make anonymous" step precedes the "analyse" step (**"first make anonymous, then analyse"** approach). The "first make anonymous, then analyse" approach rules out the "first analyse, then make anonymous" approach of statistics laws, because the value of the underlying mobile location data is lost due to applying the "make anonymous" step before the "analysis" step. In such case, the NSI would most probably not be allowed to use anonymous mobile location data for producing

[258] *Op. cit.*, F. Bieker et al. A Process for Data Protection Impact Assessment. 2016, p 29.
[259] See: Section 7.3.2 above.

official statistics under statistics laws because it will provide inaccurate results and thus not satisfy the statistical quality principles.

The Solution offers a way of reconciling the two approaches – today's new frontier technologies enable conducting the "make anonymous" and "analyse" functions in a single step without compromising privacy and confidentiality. It involves secure computations on mobile location data while maintaining the anonymity of Subscribers, i.e. making anonymous during analysis. This is achieved mainly thanks to the Intel® SGX TEE and Sharemind HI technologies underlying the Solution, complemented by a processing logic designed to return anonymous data in output (based on any combination of well-known anonymization techniques such as aggregation, randomization, generalization and suppression). We shall refer to this hybrid approach as the "**2-in-1**" approach.

8.1.2. Structure of the legal analysis

In order to assess the legal validity of the 2-in-1 approach proposed in the Sample DPIA as a result of the Project, several issues were identified during the preparation phase that need to be further analysed. Based on the central theme of the issues, the questions for legal analysis were divided into groups.

i. "Made anonymous" requirement

The first question that needs to be answered is whether further processing pseudonymous mobile location data by means of the Solution for producing official statistics qualifies as "made anonymous", in order to meet the main requirement of further processing set out in ePD Art 9(1). In order to answer it, several sub-questions can be distinguished:

1. Does the concept of "made anonymous" have its own independent meaning in EU law, which must be interpreted in a manner which fully reflects the objective of the ePD?
2. What are the criteria for deciding if further processing pseudonymous mobile location data by means of the Solution for producing official statistics qualifies as "made anonymous"?
3. Does further processing pseudonymous mobile location data by means of the Solution for producing official statistics meet those criteria?

If the answer to the first question is "no", then the further processing pseudonymous mobile location data by means of the Solution for producing official statistics cannot be in compliance with ePD Art 9, meaning that the MNO is not allowed according to *de lege lata* to make this data available to the NSI by means of the Solution. In such case, the Solution cannot be legally implemented in practice in order to process mobile location data for statistical purposes and the remainder of the questions laid out below are futile. Either the ePD Art 9(1) needs to be changed or a new legal act needs to be adopted to allow this.

If the answer to the first question is "yes", then the second question is whether further processing pseudonymous mobile location data by means of the Solution for producing official statistics is carried out by persons under the authority of the MNO,

in order to meet the additional requirement of further processing set out in ePD Art 9(3).

If the second question is also answered as “yes”, then further processing pseudonymous mobile location data by means of the Solution for producing official statistics is in compliance with ePD Art 9, meaning that the MNO is allowed according to *de lege lata* to make this data available to the NSI by means of the Solution, provided that there is a proper legal basis for it under GDPR Art 6. However, this does not assure that further processing pseudonymous mobile location data by means of the Solution for producing official statistics is a compatible further use and has a suitable legal basis nor does it provide an answer regarding who acts as (joint) controller or processor, which are critical issues for the Sample DPIA to be completed.

ii. Compatibility test

The third question is whether further processing pseudonymous mobile location data by means of the Solution for producing official statistics is compatible further use in terms of GDPR Art 6(4) and Art 5(1)(b)? In order to answer it, several sub-questions can be distinguished:

1. What is the purpose of further processing pseudonymous mobile location data by means of the Solution for producing official statistics in the context of the Sample Use Case?
2. Is this purpose compatible with the initial purposes for which the mobile location data were collected?
    a. Does the presumption of compatibility apply?
        i. If yes, are the implemented safeguards appropriate?
        ii. If no, does the processing meet the requirements of the full compatibility test?

iii. Controllership

The fourth question is who determines the purposes and means of the processing of personal data, i.e. who is the controller in terms of GDPR Art 4 (7) when further processing pseudonymous mobile location data by means of the Solution for producing official statistics?

iv. Lawfulness

The fifth question is which norm is the most suitable legal basis for further processing pseudonymous mobile location data by means of the Solution for producing official statistics? In order to answer it, several sub-questions can be distinguished:

1. Could the processing rely on consent as the legal basis provided in GDPR Art 6(1)(a)?
2. Could the processing rely on a legal obligation as the legal basis established in EU or Member State law in accordance with GDPR Art 6(1)(c) and GDPR Art 6(3)?

3. Could the processing rely on public interest/official authority as the legal basis established in EU or Member State law in accordance with GDPR Art 6(1)(e) and GDPR Art 6(3)?
4. Could the processing rely on legitimate interests as the legal basis provided in GDPR Art 6(1)(f), unless the processing is carried out by public authorities (such as Eurostat and/or national statistical institutes) in the performance of their tasks (GDPR Art 6(1) last sentence)?
5. Should there be a different legal basis, depending on whether further processing pseudonymous mobile location data by means of the Solution for producing official statistics is carried out in the proof-of-concept, pilot project or production stage?

8.2. "Made anonymous" requirement

8.2.1. Independent meaning of the concept "made anonymous"

In order to assess whether processing of mobile location data by means of the Solution qualifies as "made anonymous" under ePD Art 9(1), there is a need to understand what does the concept "made anonymous" mean under EU law.

ePD Art 9(1) reads as follows:

> *1. Where location data other than traffic data, relating to users or subscribers of public communications networks or publicly available electronic communications services, can be processed, such data may only be processed when they are made anonymous* [emphasis added]*, or with the consent of the users or subscribers to the extent and for the duration necessary for the provision of a value added service.* […]

ePD Art 2 first paragraph states that unless provided otherwise, the definitions in the DPD and in Directive 2002/21/EC of the European Parliament and of the Council of 7 March 2002 on a common regulatory framework for electronic communications networks and services (Framework Directive)[260] shall apply. DPD was repealed with effect from 25 May 2018 and references to the repealed DPD shall be construed as references to the GDPR.[261] Therefore, the reference to DPD in ePD Art 2 should be read as a reference to GDPR with regard to definitions. This means that the reference to "made anonymous" should be construed as reference to GDPR Rec 26, which is the only clause in GDPR that makes any mention of the term "anonymous".

When analysing the term "anonymous" in the context of GDPR, note should be taken of GDPR Rec 15, which establishes the principle of technological neutrality:

[260] No longer in force. Replaced by Directive (EU) 2018/1972 of the European Parliament and of the Council of 11 December 2018 establishing the European Electronic Communications Code (Recast). Text with EEA relevance. PE/52/2018/REV/1. OJ L 321, 17.12.2018, p. 36–214 (BG, ES, CS, DA, DE, ET, EL, EN, FR, GA, HR, IT, LV, LT, HU, MT, NL, PL, PT, RO, SK, SL, FI, SV). – Internet: https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A32018L1972 (10.05.2021).
[261] GDPR Art 94(1) and 94(2).

(15) In order to prevent creating a serious risk of circumvention, the protection of natural persons should be technologically neutral and should not depend on the techniques used. […]

In this light, it can be asked if "made anonymous" in ePD Art 9(1) terms or "rendered anonymous" in GDPR Rec 26 terms refers exclusively to traditional anonymisation techniques – data randomization, aggregation, suppression and generalization – or if a broader set of techniques can be inferred. For example, WP29 accepted "anonymisation" and "aggregation" as separate means of rendering anonymous in the pre-GDPR era.[262] This implies that the WP29 considered "anonymisation" and "aggregation" as the only viable options for achieving anonymity. However, as new privacy-enhancing technologies and privacy-conscientious processing models have emerged over time, the technological neutrality principle may require the data protection authorities and regulators to update their current approach to anonymity. For example, besides the two well-known techniques to anonymisation – noise addition at the input level (anonymised database) and at the output level (anonymised query result) – it is less known that "making anonymous" can also be achieved by means of anonymous processing. In that case, there is no suppression or noise needed – the underlying data remains intact and the "making anonymous" is conducted at the processing level, not only to the data.[263] The Solution implements the latter technique, i.e. anonymous processing by means of privacy-enhancing technologies, in combination with traditional anonymisation techniques and complemented by other technical, legal and organisational protection measures, as will be detailed in the following analysis (see Section 8.2.4 below) The effects of the technological neutrality principle will not be addressed in detail in this document, due to limited scope, the matter is highlighted only to point out the need for further research in this area.

Nevertheless, it should be analysed if the concept of "made anonymous" in ePD Art 9(1) terms or "rendered anonymous" in GDPR Rec 26 terms has its own independent meaning in EU law. The European Court of Justice ("**ECJ**") analysed the required level of harmonisation of national laws under the DPD, concluding that the harmonisation must be generally complete to ensure an equivalent level of protection of the rights and freedoms of individuals with regard to processing of personal data in all Member States and to ensure a high level of protection in the EU.[264] Although the objectives and principles of DPD remain sound also under GDPR, the DPD did not prevent "fragmentation in the implementation of data protection across the Union, legal uncertainty or a widespread public perception that there are significant risks to the protection of natural persons, in particular with regard to online activity".[265] Hence, the GDPR aims to improve the level of protection

[262] *Op cit.*, Article 29 Data Protection Working Party. Opinion 03/2013 on purpose limitation, 2013, p 29, p 30, p 49.
[263] D. Bogdanov, T. Siil. Anonymisation 2.0: Sharemind as a Tool for De-Identifying Personal Data - Part 2: Sharemind and anonymization. – Internet: https://sharemind.cyber.ee/anonymisation-2_0-part-2-sharemind/ (25.08.2021).
[264] European Court of Justice. Judgment of the Court. 16 December 2008, In Case C-524/06, Heinz Huber v Bundesrepublik Deutschland, sec 50, sec 51. – Internet: https://curia.europa.eu/juris/document/document.jsf?text=&docid=76077&pageIndex=0&doclang=EN&mode=lst&dir=&occ=first&part=1&cid=13821220 (03.05.2021).
[265] GDPR Rec 9.

in the EU: “In order to ensure a consistent and high level of protection of natural persons and to remove the obstacles to flows of personal data within the Union, the level of protection of the rights and freedoms of natural persons with regard to the processing of such data should be equivalent in all Member States. Consistent and homogenous application of the rules for the protection of the fundamental rights and freedoms of natural persons with regard to the processing of personal data should be ensured throughout the Union.”[266]

Similarly to DPD, the ePD also “harmonises the provisions of the Member States required to ensure an equivalent level of protection of fundamental rights and freedoms, and in particular the right to privacy, with respect to the processing of personal data in the electronic communication sector and to ensure the free movement of such data and of electronic communication equipment and services in the Community.”[267] The provisions of ePD are meant to particularise and complement GDPR for these purposes.[268] Therefore, the reference to “make anonymous” in ePD Art 9(1) should be interpreted consistently with the reference to “render anonymous” in GDPR Rec 26 – they cannot have a meaning which varies between the Member States because they define the point when data protection law ceases to apply. In conclusion, these concepts should be treated as one, having its own independent meaning in EU law, which must be interpreted in a manner which fully reflects the objectives of GDPR.[269]

Even though this topic admittedly requires further research, which is not the object of this document, it is assumed that the concepts “make anonymous” and “render anonymous” are synonyms, having an identical and independent meaning in EU law, the contents of which shall be investigated further below.

8.2.2. A potential new approach to anonymity under the GDPR

Since ePD does not provide the criteria of “made anonymous”, but refers to the GDPR instead, the next question is to clarify what are the criteria for deciding if further processing pseudonymous mobile location data by means of the Solution for producing official statistics qualifies as “made anonymous”.

GDPR talks of anonymity in GDPR Rec 26:

> *(26) The principles of data protection should apply to any information concerning an identified or identifiable natural person. Personal data which have undergone pseudonymisation, which could be attributed to a natural person by the use of additional information should be considered to be information on an identifiable natural person. To determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly. To ascertain*

[266] GDPR Rec 10.
[267] ePD Art 1(1).
[268] ePD Art 1(2).
[269] Based on analogous analysis regarding the concept of “necessity” in *Op. cit*., European Court of Justice. Judgment of the Court. 16 December 2008, In Case C-524/06, Heinz Huber v Bundesrepublik Deutschland, sec 52.

> *whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments. The principles of data protection should therefore not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable* [emphasis added]*. This Regulation does not therefore concern the processing of such anonymous information, including for statistical or research purposes.*

In broad terms, GDPR Rec 26 follows the language of DPD Rec 26:

> *(26) Whereas the principles of protection must apply to any information concerning an identified or identifiable person; whereas, to determine whether a person is identifiable, account should be taken of all the means likely reasonably to be used either by the controller or by any other person to identify the said person; whereas the principles of protection shall not apply to data rendered anonymous in such a way that the data subject is no longer identifiable* [emphasis added]*; whereas codes of conduct within the meaning of Article 27 may be a useful instrument for providing guidance as to the ways in which data may be rendered anonymous and retained in a form in which identification of the data subject is no longer possible;*

Compared to the earlier DPD Rec 26, the GDPR Rec 26 has introduced a slight change in its approach to anonymity. DPD Rec 26 explicitly considered anonymous only "data rendered anonymous in such a way that the data subject is no longer identifiable", suggesting the irreversibility of the process. In comparison, GDPR Rec 26 differentiates between two types of anonymous information:

- "information which does not relate to an identified or identifiable natural person",
- "personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable".

More interestingly, the second type – "personal data rendered anonymous" – includes two alternatives depending on the result of the process of "rendering anonymous":

1. **data subject is not identifiable** – indicating that the data subject is not identifiable in the present case, but may be identifiable if circumstances were different,
2. **data subject is no longer identifiable** – indicating that the data subject was identifiable in the past but is not identifiable at the present case or beyond, suggesting irreversibility of the process of "rendering anonymous".

This change in approach to anonymity may prove to be critical in interpreting the criteria of considering data as anonymous for the purposes of the Project. For example, if the change was made deliberately to leave room for technological development introducing new kinds of anonymisation techniques, then technologies such as TEE – and the enclaves in the Solution, more specifically – along with

relevant other technical and organisational measures may be qualified as a means to ensure that "data subject is not identifiable", i.e. the data is anonymous. If, to the contrary, the change was only accidental or unintended, it requires further analysis whether it can be interpreted expansively, so as to allow new technical and organisational measures to be classified as a means to render personal data anonymous.

It is not within the scope of this legal analysis to investigate the legislative history of this change in the approach to anonymity, as introduced by the GDPR. However, the issue merits further attention in the future as the legislative history may provide further light on the motivation of the legislators to introduce the change in approach. In the meanwhile, it is hoped that the EDPB will provide more clarity on the interpretation of the new approach in its forthcoming guidelines on anonymisation.[270]

For the purposes of the rest of the legal analysis, it is assumed that it was the intention of the legislators to change the approach to anonymity compared to DPD Rec 26 (see the next Section below). This means that any earlier interpretations and guidelines related to anonymisation, which were adopted either by the WP29, DPAs or the courts under the DPD Rec 26 may need to be reviewed and updated in light of the GDPR Rec 26. The exact scope and meaning of this change remains subject for subsequent work and is open for discussion, especially in light of the forthcoming EDPB guidelines on anonymisation.

8.2.3. Alternatively, pre-existing approach to anonymity under the DPD

i. Database sanitization paradigm

Even if it is concluded that the different wording of GDPR Rec 26 was not meant to change the approach to anonymity compared to DPD Rec 26, the pre-existing interpretations of DPD Rec 26 and the regulatory guidelines related thereto apply to interpreting GDPR Rec 26, as well. For example, WP29 issued the Opinion 05/2014 on Anonymisation Techniques ("**WP29 Opinion on Anonymisation Techniques**") in 2014[271], which the EDPB has not endorsed[272] but which has been repeatedly referred to in the opinions of EDPS and EDPB and thus remains relevant also in interpreting GDPR, at least until a new guideline is provided on the matter by the EDPB.

In its Opinion on Anonymisation, the WP29 analysed the effectiveness and limits of existing anonymisation techniques against the EU legal background of data protection and provided recommendations for a cautious and responsible use of

[270] European Data Protection Board. EDPB Work Programme 2021/2022, p 4. – Internet: https://edpb.europa.eu/system/files/2021-03/edpb_workprogramme_2021-2022_en.pdf (30.04.2021).
[271] *Op. cit.*, Article 29 Data Protection Working Party, Opinion 05/2014 on Anonymisation Techniques, 2014.
[272] European Data Protection Board. GDPR: Guidelines, Recommendations, Best Practices. Endorsement of GDPR WP29 Documents. – In the Internet: https://edpb.europa.eu/our-work-tools/general-guidance/gdpr-guidelines-recommendations-best-practices_en (21.01.2021).

these techniques to build a process of anonymisation.[273] According to the analysis of WP29, for data to be anonymised it must be stripped of sufficient elements such that the data subject can no longer be identified by either the controller or a third party and such process is irreversible. [274]

The approach to anonymisation techniques applied in the WP29 Opinion on Anonymisation has been criticized as too narrow in scope because it concentrates on database sanitization techniques.[275] "Examples of such sanitization techniques are mechanisms that rely on data suppression and generalization (known as anonymization techniques) […], and those that rely on noise addition like in differential privacy […]. Nevertheless, there is neither well-defined scheme to evaluate the robustness of sanitization techniques, nor a clear understanding for "when data is regarded as well-sanitized".".[276]

The limitations of database sanitization techniques are well known among the privacy engineering community. WP29 also acknowledges their shortcomings.[277] In practice, several other techniques are emerging that can be used to de-identify personal data or otherwise disguise the identity (e.g. secret sharing, multi-party computation, (fully) homomorphic encryption and TEE). However, such techniques have not been considered or not even mentioned in the WP29 Opinion on Anonymisation Techniques.

WP29 has, at least in once instance, acknowledged that "[d]isguising identities can also be done in a way that no re-identification is possible, e.g. by one-way cryptography, which creates in general anonymised data."[278] On another instance in the same document, WP29 concluded that "re-identification of the data subject may have been excluded in the design of protocols and procedure".[279] EDPB, on the other hand, has declared in its Corona App Guidelines that many options exist for effective anonymisation but data cannot be anonymised on their own – "only datasets as a whole may or may not be made anonymous. […] any intervention on a single data pattern (by means of encryption, or any other mathematical

[273] *Op. cit.,* Article 29 Data Protection Working Party, Opinion 05/2014 on Anonymisation Techniques, 2014, p 5.
[274] *Ibid.*, p 5.
[275] S. Schiffner, B. Berendt, et al. Towards a Roadmap for Privacy Technologies and the General Data Protection Regulation: A transatlantic initiative. In: Medina M., Mitrakas A., Rannenberg K., Schweighofer E., Tsouroulas N. (eds). Privacy Technologies and Policy. APF 2018. Lecture Notes in Computer Science, vol 11079. Springer, Cham., 2018, p 9. – Internet: https://doi.org/10.1007/978-3-030-02547-2_2 (01.05.2021).
(01.05.2021).
[276] A. Kassem, G. Acs, C. Castelluccia. Differential Inference Testing A Practical Approach to Evaluate Anonymized Data. [Research Report] INRIA. 2018, p 2. - Available in the Internet at: https://hal.inria.fr/hal-01681014/ (01.05.2021).
[277] *Op. cit.*, Article 29 Data Protection Working Party, Opinion 05/2014 on Anonymisation Techniques, 2014, p 10.
[278] Article 29 Data Protection Working Party. Opinion 4/2007 on the concept of personal data. Adopted on 20th June. 01248/07/EN, WP 136, p 18. – Internet: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf (01.05.2021).
[279] *Ibid., p 20.*

transformation) can at best be considered a pseudonymisation."[280] In the same document, EDPB has also referred to a paper introducing four models for the privacy-conscientious use of mobile phone data, which are seen by the authors of that paper to overcome the limits of traditional data anonymisation methods.[281] It is hoped that the forthcoming EDPB guidelines on anonymisation[282] will take a closer look at the emerging privacy-enhancing technologies and qualify them as state-of-the-art anonymisation techniques.

It is clear that the Solution as a whole and the TEE in the form of the enclaves within the Solution do not fit well into the data sanitization paradigm that the WP29 relied on when adopting its Opinion on Anonymisation Techniques. They simply do not concern with a data analyst directly manipulating a database – all the calculations on mobile location data are run automatically in the Solution in a closed environment where nobody has access to them. The Sample Use Case does include an SDC step, where data sanitization techniques are applied but it comes only after the statistical analysis has been conducted and the resulting aggregate data is further processed to meet the statistical confidentiality criteria. Even then, the SDC are applied in complete isolation from the data analyst – no person or organisation is directly involved in all the intermediate processing steps, nor can they see the different forms and calculations of data.

ii. Hybrid Anonymisation Paradigm

a) The contextual nature of anonymisation

New privacy-enhancing technologies have emerged since the adoption of the WP29 Opinion on Anonymisation Techniques. As the EDPB is in the process of preparing a new guideline on anonymisation, it merits analysis whether the Solution meets the criteria of anonymisation beyond the data sanitization paradigm – could the Solution be qualified as a process of anonymisation under the the GDPR Rec 26, if EDPB were to accept a more liberal anonymisation paradigm, which recognises other anonymisation techniques besides database sanitization ("**Hybrid Anonymisation Paradigm**")? In order to do that under the presumption that the GDPR did not introduce a significant change in approach to anonymity, one would need to consider the more objective conditions for an anonymisation process, which can be relied on independent of the specific anonymisation techniques applied.

When leaving the statements specific to data sanitization techniques aside, the WP29 Opinion on Anonymisation Techniques offers valuable reference points for determining what is anonymisation:

---

[280] *Op. cit.*, European Data Protection Board. Guidelines 04/2020 on the use of location data and contact tracing tools in the context of the COVID-19 outbreak, 2020, p 6.
[281] *Ibid.*, p 6. – Referring to: Y.-A. de Montjoye, S. Gambs, V. Blondel, G. Canright, N. de Cordes, S. Deletaille, K. Engø-Monsen, M. Garcia-Herranz, J. Kendall, C. Kerry, G. Krings, E. Letouzé, M. Luengo-Oroz, N. Oliver, L. Rocher, A. Rutherford, Z. Smoreda, S. Steele, E. Wetter, Alex "Sandy" Pentland L. Bengtsson. Comment: On the privacyconscientious use of mobile phone data. Scientific Data 5 (December 2018): 180286. – Internet: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6289108/pdf/sdata2018286.pdf (02.05.2021).
[282] European Data Protection Board. EDPB Work Programme 2021/2022, p 4. – Internet: https://edpb.europa.eu/system/files/2021-03/edpb_workprogramme_2021-2022_en.pdf (30.04.2021).

1. it confirms that the DPD does not clarify how an anonymisation process should or could be performed[283]. Instead, the focus is on the outcome: that data should be such as not to allow the data subject to be identified via "all" "likely" and "reasonable" means;[284]
2. it highlights that the anonymisation process must be irreversible, defining anonymisation as "a technique applied to personal data in order to achieve irreversible de-identification" – note, however, that the requirement of irreversibility becomes questionable if it is confirmed that the approach to anonymity has changed. Under DPD Rec 26, the data subject was supposed to be "no longer identifiable" as a result of anonymisation, referring to irreversibility of the process. GDPR Rec 26 seems to allow the data subject to remain non-identifiable at present circumstances, while not expressly ruling out the possibility that data subject may be identifiable in other circumstances;[285]
3. it points out that there is an "inherent residual risk of re-identification linked to any technical-organisational measure aimed at rendering data "anonymous".";[286]
4. it clarifies that since research, tools and computational power evolve, it is neither possible nor useful to provide an exhaustive enumeration of circumstances when identification is no longer possible.[287]

WP29 Opinion 4/2007 on the concept of personal data[288] ("**WP29 Opinion on the Concept of Personal Data**") also emphasizes that qualifying a process as anonymisation is context-specific: "the assessment whether the data allow identification of an individual, and whether the information can be considered as anonymous or not depends on the circumstances, and a case-by-case analysis should be carried out with particular reference to the extent that the means are likely reasonably to be used for identification as described in Recital 26."[289]

b) Identifiability test

Based on the above, there is no single technique that can be qualified as anonymisation in terms of GDPR or DPD under all circumstances. Each technique needs to be evaluated based on the particular context of a specific case. However, data protection authorities have provided some guidance on how to determine anonymity. Based on that, it seems that the general criteria for assessing the robustness of the anonymisation process have remained stable under the GDPR, irrespective of whether the approach to anonymity has changed from DPD to GDPR.

EDPB has recently summarised its take on identifiability in the context of health research in its Document on Health Research. According to EDPB, "[t]he

[283] *Op. cit.,* Article 29 Data Protection Working Party, Opinion 05/2014 on Anonymisation Techniques, 2014, p 5.
[284] *Ibid.*
[285] *Ibid.*, p 5, 7.
[286] *Ibid.*, p 7.
[287] *Ibid.*, p 8.
[288] *Op. cit.*, Article 29 Data Protection Working Party. Opinion 4/2007 on the concept of personal data, 2007.
[289] *Ibid.*, p 21.

determination of whether information is anonymous must be made by application of the test of identifiability outlined in Recital 26 GDPR"[290] ("**identifiability test"**):

> […] *To determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly. To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments.* […]

EDPB further specified what should be taken into account when making an assessment as to the reasonable likelihood of identifiability:

1. all the factors outlined in Recital 26 of GDPR must be considered,[291]
2. the WP29 Opinion on Anonymisation Techniques should be taken into account,[292]
3. "[a]ny such assessment should be made along the lines suggested by the CJEU in *Breyer*, which refers to Recital 26 of Directive 95/46/EC, looking at the legal and practical means by which re-identification may be effected by the use of additional data in the hands of third parties."[293]
4. ongoing advancements in available technological means and progress made in the field of re-identification.[294]

These considerations will be addressed one by one below as follows, with due consideration of the WP29 Opinion on Anonymisation Techniques where relevant.

1) Key risk factors

In order to identify "all objective factors" required under GDPR Rec 26, different guidelines of the data protection authorities in the EU are instructive.

In the pre-GDPR era, the WP29 provided in its Opinion on Anonymisation Techniques a non-exhaustive list of key risk factors to be taken into consideration when evaluating the potential identifiability of a given dataset that undergoes anonymisation according to the different available techniques[295]:

- **means to reverse anonymisation** – "data controllers should focus on the concrete means that would be necessary to reverse the anonymisation

[290] European Data Protection Board. EDPB Document on response to the request from the European Commission for clarifications on the consistent application of the GDPR, focusing on health research. Adopted on 2 February 2021, p 11, sec 45. – Internet: https://edpb.europa.eu/our-work-tools/our-documents/other-guidance/edpb-document-response-request-european-commission_en (07.05.2021).
[291] *Ibid.*, p 11, sec 46.
[292] *Ibid*.
[293] *Ibid*.
[294] *Ibid.*, sec 47.
[295] *Op. cit.,* Article 29 Data Protection Working Party, Opinion 05/2014 on Anonymisation Techniques, 2014, p 10.

technique, notably regarding the cost and the know-how needed to implement those means and the assessment of their likelihood and severity".[296]

   a. For example: the anonymisation effort and costs (time and resources) should be balanced against the increasing low-cost availability of technical means to identify individuals in datasets, the increasing public availability of other datasets (e.g. as a result of open data policies), and the many examples of incomplete anonymisation entailing subsequent adverse, sometimes irreparable effects on data subject.[297]

- **individual event aggregation** – "when a data controller does not delete the original (identifiable) data at event-level, and […] hands over part of this dataset (for example after removal or masking of identifiable data), the resulting dataset is still personal data. Only if the data controller would aggregate the data to a level where the individual events are no longer identifiable, the resulting dataset can be qualified as anonymous."[298]
   a. "For example: if an organisation collects data on individual travel movement, the individual travel patterns at event level would still qualify as personal data for any party, as long as the data controller (or any other party) still has access to the original raw data, even if direct identifiers have been removed from the set provided to third parties. But if the data controller would delete the raw data, and only provide aggregate statistics to third parties on a high level, such as 'on Mondays on trajectory X there are 160% more passengers than on Tuesdays', that would qualify as anonymous data."[299]
- **potential identifiability by singling out, linkability and inference** – "removing directly identifying elements in itself is not enough to ensure that identification of the data subject is no longer possible. It will often be necessary to take additional measures to prevent identification, […] depending on the context and purposes of the processing for which the anonymised data are intended."[300]

In the pre-GDPR era, the WP29 often referred to the assessment of robustness of the chosen anonymisation technique as a separate issue to be addressed in the context of identifiability. According to WP29, the DPD suggested a "means … reasonably to be used" criterion for assessing whether the anonymisation process is sufficiently robust, i.e. whether identification has become "reasonably" impossible.[301] This criterion is embodied in DPD Rec 26: "whereas, to determine whether a person is identifiable, account should be taken of <u>all the means likely reasonably to be used [emphasis added]</u> either by the controller or by any other person to identify the said person".

[296] *Ibid.*, pp 8-9.
[297] *Ibid.*, p 9.
[298] *Ibid.*
[299] *Ibid.*
[300] *Ibid.*
[301] *Op. cit.*, Article 29 Data Protection Working Party, Opinion 05/2014 on Anonymisation Techniques, 2014, p 8.

The criterion survived also in the GDPR Rec 26, albeit with losing "all" in front of the "means", a slight reordering of words ("reasonably likely" instead of "likely reasonably") and a specification of "all objective factors" to be taken into account:

> *[…] To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments. […]*[302]

The sample list of objective factors expressly listed in GDPR Rec 26 are not new, *per se*. The requirement to take into account "all the factors at stake" was introduced already in the WP29 Opinion on the Concept of Personal Data, where the "means to identify" in the context of DPD Rec 26 where further clarified.[303] WP29 began the analysis by stating that a mere hypothetical possibility to single out the individual is not enough to consider the person as "identifiable" – if that possibility does not exist or is negligible, taking into account "all the means likely reasonably…" criterion, the person should not be considered as "identifiable" and the information would not be considered as "personal data".[304] It continued by requiring that "all the means likely reasonably…" criterion should in particular take into account all the factors at stake, for example:

- the cost of conducting identification,
- the intended purpose of the data processing (whether identification of data subjects is embedded in the purposes and means of the processing),
- the way the processing is structured,
- the advantage expected by the controller,
- the interests at stake for the individuals,
- the risk of organizational dysfunctions (e.g. breaches of confidentiality duties),
- the risk of technical failures,
- the time period during which data are stored,
- the state of the art in technology at the time of processing and the possibilities for development during the period for which the data will be processed.[305]

In addition, WP29 Opinion on the Concept of Personal Data can be instructive when considering if and how the processing contributes to identifiability of an individual data subject. It offers three alternative elements in determining whether information "relates to" an individual and thus makes the person indirectly identifiable in terms of GDPR:

- **content** – "information is given about a particular person, regardless of any purpose on the side of the data controller or of a third party, or the impact of that information on the data subject. […] this has to be assessed in the light of all circumstances surrounding the case."[306]

---

[302] GDPR Rec 26.
[303] *Op. cit.*, Article 29 Data Protection Working Party. Opinion 4/2007 on the concept of personal data, 2007, p 15.
[304] *Ibid.*
[305] *Ibid.*, pp 15-16.
[306] *Ibid.*, p 10.

- **purpose** – "when the data are used or are likely to be used, taking into account all the circumstances surrounding the precise case, with the purpose to evaluate, treat in a certain way or influence the status or behaviour of an individual."[307]
- **result** – "Despite the absence of a "content" or "purpose" element, data can be considered to "relate" to an individual because their use is likely to have an impact [emphasis added] on a certain person's rights and interests, taking into account all the circumstances surrounding the precise case. […] it is not necessary that the potential result be a major impact. It is sufficient if the individual may be treated differently from other persons as a result of the processing of such data."[308]

The EDPB has recently emphasized in its Corona App Guidelines that data is anonymised if it passes the "reasonability test" ("**reasonability test**"):

15 *Anonymisation refers to the use of a set of techniques in order to remove the ability to link the data with an identified or identifiable natural person against any "reasonable" effort. This "reasonability test" must take into account both objective aspects (time, technical means) and contextual elements that may vary case by case (rarity of a phenomenon including population density, nature and volume of data). If the data fails to pass this test, then it has not been anonymised and therefore remains in the scope of the GDPR.*[309]

EDPB bases the robustness of anonymisation on three criteria:

1. **singling-out** – isolating an individual in a larger group based on the data;
2. **linkability** – linking together two records concerning the same individual;
3. **inference** – deducing, with significant probability, unknown information about an individual.[310]

2) Legal and practical means of identification (the case of Breyer)

When conducting the reasonability test as part of the identifiability test, consideration should be taken of the fact that there is an ongoing dispute, mainly in German academic writings and case-law, whether an absolute (or objective) or a relative (or subjective) approach has to be used for the assessment of the identifiability of data subjects.[311] To summarize the dispute:

- according to the **absolute (or objective) approach**, data may be regarded as being personal data even if only a third party is able to determine the identity

[307] *Ibid.*
[308] *Ibid.*, p 11.
[309] *Op. cit.*, European Data Protection Board. Guidelines 04/2020 on the use of location data and contact tracing tools in the context of the COVID-19 outbreak, 2020, p 5.
[310] *Ibid.*
[311] G. Spindler, P. Schmechel. Personal Data and Encryption in the European General Data Protection Regulation. 7 (2016) JIPITEC 163 para 1, p 1. – Internet: https://www.jipitec.eu/issues/jipitec-7-2-2016/4440 (02.05.2021).

of the data subject.[312] This approach is represented in the national law of a minority of EU Member States, such as France.[313]

- In contrast, the **relative (or subjective) approach** regards data as personal data in relation to an entity because they allow the user to be precisely identified.[314] Therefore, only realistic chances of combining data in order to identify an individual are taken into account.[315] This approach is presented in the national law of countries such as Ireland and Germany.[316] In other words, "if an unreasonable effort were required to reidentify anonymized data, then it would no longer be personal data."[317]

GDPR Rec 26 seems to have both "absolute" and "relative" elements. For example, on the one hand, the reference to "means reasonably likely to be used […] by another person" veers towards an absolute approach, because this third person could be any person in the world. On the other hand, the term "means reasonably likely to be used" suggests limitations through relative elements, in particular the notion of "reasonably". The objective factors for interpreting the "means reasonably likely to be used" illustrate a further attempt to weigh towards relative approach. [318]

The ECJ has addressed the matter of absolute vs relative approach to identifiability in its 19.10.2016 judgment in the Patrick Breyer v Bundesrepublic Deutschland case.[319] With regard to the means likely reasonably to be used by both the controller and by "any other person" under DPD Rec 26, the ECJ concluded that this wording suggests "it is not required that all the information enabling the identification of the data subject must be in the hands of one person" for information to be treated as "personal data" within the meaning of DPD.[320] In other words, different elements of personal data can be held by different persons – this, in itself, does not render the personal data anonymous. It seems the ECJ does not clearly endorse the absolute approach with this interpretation, but rather leans towards the relative approach.

---

[312] European Court of Justice. Judgment of the Court (Second Chamber), 19 October 2016, C-582/14, Patrick Breyer v Bundesrepublic Deutschland, sec 25. – Internet: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A62014CJ0582&qid=1619966262083 (02.05.2021).
[313] J. Scheibner, J.L. Raisaro, J.R. Troncoso-Pastoriza, M. Ienca, J. Fellay, E. Vayena, J. Hubaux. Revolutionizing Medical Data Sharing Using Advanced Privacy-Enhancing Technologies: Technical, Legal, and Ethical Synthesis. Journal of Medical Internet Research, Vol 23, No 2 (2021): February, p 6. – Internet: https://www.jmir.org/2021/2/e25120/ (02.05.2021). – Referring to: M. Finck, F. Pallas. They who must not be identified - distinguishing personal from non-personal data under the GDPR. International Data Privacy Law, Volume 10, Issue 1, February 2020. – Internet: https://academic.oup.com/idpl/article/10/1/11/5802594?login=true (02.05.2021).
[314] *Op. cit.*, European Court of Justice. Judgment of the Court (Second Chamber), 19 October 2016, C-582/14, Patrick Breyer v Bundesrepublic Deutschland, sec 25.
[315] *Op. cit.*, G. Spindler, P. Schmechel. Personal Data and Encryption in the European General Data Protection Regulation, 2016, pp 165-167.
[316] *Op. cit.*, J. Scheibner, et al. Revolutionizing Medical Data Sharing Using Advanced Privacy-Enhancing Technologies: Technical, Legal, and Ethical Synthesis, 2021, p 6.
[317] *Ibid.*, Referring to: G. Spindler, P. Schmechel. Personal Data and Encryption in the European General Data Protection Regulation, 2016.
[318] *Op. cit.*, G. Spindler, P. Schmechel. Personal Data and Encryption in the European General Data Protection Regulation, 2016, p 166.
[319] *Op. cit.*, European Court of Justice. Judgment of the Court (Second Chamber), 19 October 2016, C-582/14, Patrick Breyer v Bundesrepublic Deutschland.
[320] *Ibid.*, sec 43.

However, in the subsequent sections, the ECJ appears to take a turn towards the absolute approach, relying on the Advocate General's opinion in the same case. According to the ECJ, what needs to be determined is whether the possibility to combine the personal data elements held by different persons constitutes a means likely reasonably to be used to identify the data subject.[321] "[T]hat would not be the case if the identification of the data subject was prohibited by law or practically impossible on account of the fact that it requires a disproportionate effort in terms of time, cost and man-power, so that the risk of identification appears in reality to be insignificant."[322]

The Advocate General's opinion underlying the final judgment has been interpreted as "a vote for a rather *absolute* approach"[323]. The Advocate General's interpretation allowed "even the *possibility* of obtaining the data" in order it to be considered identifiable, which can be a significant limitation of the above mentioned relative elements and widens the material scope of the DPD and, consequently, also that of the GDPR significantly.[324] Such further broadening can be found in the Advocate General's statement that alone the sheer *potential* possibility of identification shall be sufficient, not even requiring that the relevant party actually receives the identifying information.[325]

The ECJ seems to have confirmed the Advocate General's opinion by using the phrase "prohibited by law or practically impossible". To further cement its position, the ECJ explained that even though the national law did not allow the missing personal data element to be shared with third parties in order to identify the data subject under normal circumstances, there is an alternative legal route to obtaining this information in the event of cyber attacks, so that necessary steps can be taken to obtain that information and to bring criminal proceedings. The ECJ concluded, that this qualifies as "the means which may likely reasonably be used to identify the data subject, with the assistance of other persons […]".[326]

In conclusion, the ECJ endorsed legal techniques as a means to identify an individual. By doing so, the ECJ also voted for a rather absolute approach, along the lines with the Advocate General's opinion. Presumably, there is always some potential legal route to request access to information in order to initiate court proceedings. As a result, "virtually all data would have to be considered as personal data, which would, in the end, weaken the data protection framework and could

[321] *Ibid.*, sec 45.
[322] *Ibid.*, sec 46.
[323] *Op. cit.*, G. Spindler, P. Schmechel. Personal Data and Encryption in the European General Data Protection Regulation, 2016, p 168.
[324] *Ibid., p 167.* - Referring to: Opinion of Advocate General Campos Sánchez-Bordona, delivered on 12 May 2016, Case C-582/14 – Patrick Breyer v Bundesrepublik Deutschland, sec 72. – Internet: https://curia.europa.eu/juris/document/document.jsf?docid=178241&doclang=EN (02.05.2021).
[325] *Ibid.* – Referring to: Opinion of Advocate General Campos Sánchez-Bordona, delivered on 12 May 2016, Case C-582/14 – Patrick Breyer v Bundesrepublik Deutschland, sec 77. – Internet: https://curia.europa.eu/juris/document/document.jsf?docid=178241&doclang=EN (02.05.2021).
[326] *Op. cit.*, European Court of Justice. Judgment of the Court (Second Chamber), 19 October 2016, C-582/14, Patrick Breyer v Bundesrepublic Deutschland, sec 47-48.

make it unworkable"[327] and "could result in "perverse incentives" for controllers to abandon anonymisation and therefore increase, rather than relieve, privacy risks."[328]

The impact of the ECJ decision in the Patrick Breyer v Bundesrepublic Deutschland case can be far-reaching, if followed widely in the privacy engineering and privacy-by-design practices. It may mean that in addition to technical risk assessments, controllers will have to start conducting legal risk assessment in order to prove that there is no or close-to-minimum possibility to bring criminal or other legal proceedings for obtaining the missing elements of personal data from third parties. Were this to be the case, the risks and costs of data protection may become unbearable, especially to small and medium sized enterprises. In order not to speculate further, the development of this matter will have to followed in practice to make solid conclusions.

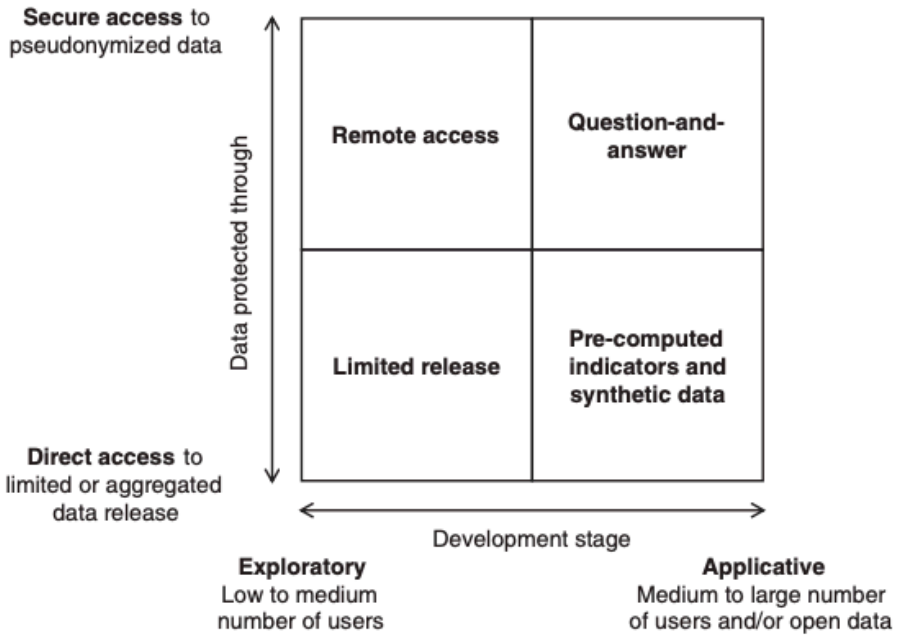
3) Available technological means and progress made in the field of re-identification

As recently acknowledged by the EDPB in its Corona App Guidelines, location data originating from telecom operators and/or information society services are known to be notoriously difficult to anonymise. For this reason, it is crucial for any controller implementing anonymisation solutions to monitor recent developments in the field of anonymisation processes and re-identification attacks, which are active fields of research. "To achieve anonymisation, location data has to be carefully processed in order to meet the reasonability test. In this sense, such a processing includes considering location datasets as a whole, as well as processing data from a reasonably large set of individuals using available robust anonymisation techniques, provided that they are adequately and effectively implemented."[329]

The EDPB refers to recent scholarship on the topic of anonymising location data in its Corona App Guidelines, according to which there are four models for the privacy-conscientious use of mobile phone data which have been "designed to fall under the legal umbrella of anonymous use of the data" [330] and provide "a reasonable balance between utility and privacy"[331] (see Figure 12 below). Although none of them is considered to be a silver bullet according to the authors, each is believed to provide a reasonable balance between utility and privacy.[332] For the purposes of demonstrating the multiple layers and dimensions of the 2-in-1 approach applied in the Solution, a brief overview of each model is provided below.

[327] *Op. cit.*, G. Spindler, P. Schmechel. Personal Data and Encryption in the European General Data Protection Regulation, 2016, p 168.
[328] *Ibid.*
[329] *Op. cit.*, European Data Protection Board. Guidelines 04/2020 on the use of location data and contact tracing tools in the context of the COVID-19 outbreak, 2020, p 6.
[330] *Op. cit.*, Y.-A. de Montjoye, et al. Comment: On the privacyconscientious use of mobile phone data. Scientific Data, 2018, p 2.
[331] *Ibid.*
[332] *Ibid*.

Figure 12 – Matrix of the four models for the privacy-conscientious use of mobile phone data[333]



Limited release – the closest to traditional sharing of data where a mobile phone dataset is transformed in-house and a copy of the data is given to third-parties under a legal contract. Because the transformed data is released directly to the users, the data controller loses technical control over the data. "This significantly increases the risk of the data to be stolen, uploaded online, or to be part of a data breach. It puts a lot of weight on the data anonymization procedure. Because of this, we consider re-identification using auxiliary location information to be the main privacy threat in the limited release model: re-identification would allow an attacker to link the released data about one to all of the users back to their identities."[334]

Pre-computed indicators and synthetic data – indicators derived from mobile phone data are released to third-parties. Synthetic data generated by the model and preserving pre-defined statistical properties of the original data can equivalently be released. "[…] the main privacy threats for pre-computed indicators and synthetic data to be questions around the notion of "group privacy", which pertains to all release types. Definitions vary but, intuitively, the idea is that one's individual privacy might be violated if information about a group he belongs to is revealed. Aggregated or anonymized data might indeed reveal sensitive information on groups and could lead to stigmatization or discrimination. In the case of mobile phone data, the privacy of a specific ethnic, or religious or minority group might, for example, be endangered if information about their behavior were to be revealed."[335]

[333] *Ibid*.
[333] European Data Protection Board. EDPB Work Programme 2021/2022, p 4. – Internet: https://edpb.europa.eu/system/files/2021-03/edpb_workprogramme_2021-2022_en.pdf (30.04.2021)
[334] *Ibid.*, p 3.
[335] *Ibid.*

Remote access – the first model using the privacy-through-security approach. "Here, the data are not released but instead stay within the premises and under the control of the operator (or an authorized entity) and are analysed remotely. The data processing takes place within the operator's premises and only aggregated data leave the secure area. […] the data controller does not have to relinquish all control over the data. The controller can supervise who accesses the data (having users registering, signing a DUA, setting restrictions on IP addresses), how the data are being used (e.g., through active monitoring of the secured environment or by controlling the output), and can ensure that no individual-level or raw data leave the server (through a manual approval process or by monitoring the amount of data leaving the server). While they do not remove all possible risks, these security-based mechanisms already strongly limit the risks of the data to be re-identified en masse and misused. This, in turn, allows the data controller to transform the data less aggressively, for instance only removing phone numbers and other direct personal identifiers, potentially along with limited temporal and spatial coarsening. This limited transformation as well as the ability to access data in near-real time strongly increase the utility and possible uses of the data."[336] "We see the main privacy threat for the remote access model to be the risk of a targeted user to be reidentified. Because the data analysis happens within a secured and controlled environment, the mass reidentification of users and exfiltration of their data is very unlikely. A secondary threat would be for the server holding the data to be compromised. While not impossible, we do not consider this risk to be significantly higher than the risk of the server currently holding the data to be compromised."[337]

Question-and-answer – "pushes the privacy through-security approach one step further: the data stays within the premises of the operator but third parties now only access the data through a question-and-answer system (e.g., SafeAnswers [footnote excluded] or SQL queries [footnotes excluded]). Questions are asked in the form of a piece of code whose answers are computed using the pseudonymized data. These are validated by the system before being sent back to the user through the API. Answers can be at the level of individuals or, more often, groups of individuals."[338] "On top of this, because the framework and language used to ask the questions as well as the user-facing API are standardized, more advanced and automated security and auditing mechanisms can be put in place [e.g. the system can validate the code, ensure the aggregation mechanism protects individuals' privacy, ensuring that a certain level of noise is added or guaranteeing differential privacy, every question asked can be fully logged etc]."[339] "Since the use of the data is tightly controlled, we consider the server being compromised to be the main privacy threat. However, as for the remote access model, we do not consider this risk to be significantly higher than the risk of any places where the data would be digitally stored (server, laptops, etc.) to be compromised. While the likelihood of an attacker being able to infer information about a specific re-identified user through the QA API is not null (these attacks served as motivation for mechanisms such as differential privacy [footnote excluded]), we consider this risk to be moderate when combined with defense-in-depth mechanisms. In both the remote access and question-and-

[336] *Ibid.*, pp 3-4.
[337] *Ibid.*
[338] *Ibid.*, p 4.
[339] *Ibid.*

answer model, the data controller does not lose technical control over the data and measures can always be taken as response to a potential privacy breach."[340]

8.2.4. Further processing pseudonymous mobile location data by means of the Solution for producing official statistics as "making anonymous"

The last step in answering the first question of the legal analysis is to assess whether further processing pseudonymous mobile location data by means of the Solution for producing official statistics meets the criteria of "made anonymous". As the WP29 Opinion on Anonymisation Techniques is out-dated due to its focus on data sanitization techniques and the new guidelines on anonymisation are under preparation at the EDPB, an effort is made to operationalise the currently available criteria described in Section 8.2.3 above by combining the identifiability test and reasonability test into a coherent framework as follows:

Figure 13 – Identification factors

Identification factors	**Technical robustness**	**Legal/Organisational robustness**
Objective aspects	- Technical means (state of the art and possibilities for development) - Risk of technical failures - Cost of identification (time and resources)	- Governance of the rights and roles - Management of shared resources - The way the processing is structured - Risk of organizational dysfunctions (e.g. breaches of confidentiality duties) - Cost of identification (time and resources)
Contextual elements	- Singling out - Linkability - Inference	- Content, purpose and result of processing - Expected advantage to the controllers vs interests of individuals - Rarity of phenomenon (including population density, nature and volume of data) - Data storage - Possibility to combine the personal data elements held by different persons

[340] *Ibid.*

Figure 13 summarizes the currently available guidance from court practice and data protection authorities with regard to assessing whether a "reasonable" effort to link the data with an identified or identifiable natural person is precluded. According to the proposed framework, the assessment is conducted in two dimensions, involving both the objective and contextual viewpoints:

- **technical robustness** – subject to the reasonability test (elaborated above in Section 8.2.3.ii.b)1));
- **legal and organisational robustness** – lawful means likely reasonably to be used to identify an individual (elaborated above in Section 8.2.3.ii.b)).

For the purposes of the Sample DPIA, it is assumed that once both criteria are met, "making anonymous" can be considered as successful. However, it is possible that the EDPB may come to a different position regarding the scope of the criteria to be considered in order to achieve anonymity. For example, it may further specify the criteria by introducing a hierarchy between them. If and how it will do so, remains to be seen in the future regulatory practice and guidelines (e.g. the forthcoming guidelines on anonymisation).

We claim that the combination of the following hybrid techniques (as will be explained below) with the relevant legal and organisational measures achieves a high level of protection which is, at the minimum, equivalent to the criteria of "making/rendering data anonymous" both in terms of the potential new approach to anonymity under GDPR Rec 26, as well as the pre-existing approach to anonymity under DPD Rec 26 (and the ePD Art 9(1)), whichever will be confirmed by relevant DPAs, EDPB, EDPS or the courts. In what follows, the re-identifiability assessment is carried out in the context of further processing pseudonymous mobile location data by means of the Solution for producing official statistics, taking the Sample Use Case as an example to illustrate the potential contextual elements that need to be considered in the pilot stage and production stage in the future.

i. Technical robustness

a) Objective aspects

1) Technical means

As a result of the Project, Cybernetica AS developed the Solution in order to enable processing of mobile location data for producing European statistics in the context of the Sample Use Case in compliance with the standards and regulations applicable both to MNO and NSI. This task required an interdisciplinary team of specialists, strong project and change management capabilities, understanding of the MNO business processes and NSI statistics production processes, and ability to resolve uncertainties at the intersection of different domains of expertise. To the best of our knowledge, the Project has introduced several novelties:

- the first time that privacy-enhancing technologies are applied in developing official statistics,
- the first time that synthetic mobile location data is processed in developing official statistics,

- the first time that official statistics are developed based on secondary use of data sources in the private sector (outside the data sources in the public administrative system),
- the first time that a statistical authority relies on the assistance of third parties to conduct statistical analysis,
- the first time that mobile location data are "made anonymous" using traditional anonymisation techniques combined with novel techniques.

Taken as a whole, no other technique or approach has managed to achieve the same results. Thus, we believe that the Project helped push forward the state-of-the art in terms of techniques ensuring against "means reasonably likely to be used to identify the natural person". For the purposes of the remainder of the legal analysis, we shall refer to these techniques as "**hybrid techniques**", so as to demarcate the lines between the emerging privacy-enhancing technologies capable of keeping a data subject un-identifiable on a permanent basis, in combination with relevant organisational and legal measures (see Section ii.b)1) below for more details on the specific protection measures), from the one side, and traditional anonymisation techniques (database sanitation), from the other side.

Based on the matrix of the privacy-conscientious use of mobile phone data introduced above, we present the Solution as a combination of key technical elements from the proposed four models:

- **limited release**: the Solution implements the "in-house transformation of dataset" element from the "limited release" model. Just as in case of the "limited release" model, the transformation is aimed at "both adding technical difficulties to attempts at re-identifying individuals and at limiting the amount of information that could be uncovered if the data were to be re-identified"[341].
- **remote access**: from the access perspective, the Solution follows the "remote access" model as the mobile location data never leaves the MNO. Furthermore, only MNO-ND is able to access raw mobile location data (as before) and MNO-VAD is only able to access periodically pseudonymised mobile location data (with short pseudonymisation periods, such as 24h in the Sample Use Case).
- **pre-computed indicators**: by employing cryptographic techniques (TEE), the Solution is able to securely link mobile location data records over several pseudonymisation periods, thus keeping most of the utility of the mobile location data. The analysis itself is pre-approved by several independent stakeholders (Auditors, Enforcers) and as such seems closest to the "pre-computed indicators" model.
- **question-and-answer**: however, since the Solution is extensible with new analysis other than the example Sample Use Case analysed here, we claim that it is actually closer to the "question-and-answer" model. The only difference is that instead of approving individual answers to be released, the analysis code itself is pre-approved. All automated controls in the "question-and-answer" model can be pre-approved in such way. Additionally, in

[341] *Op. cit.*, Y.-A. de Montjoye, et al. Comment: On the privacyconscientious use of mobile phone data. Scientific Data, 2018, p 3.

conformity with the recommendation tied to the “question-and-answer” model[342], the Solution includes:

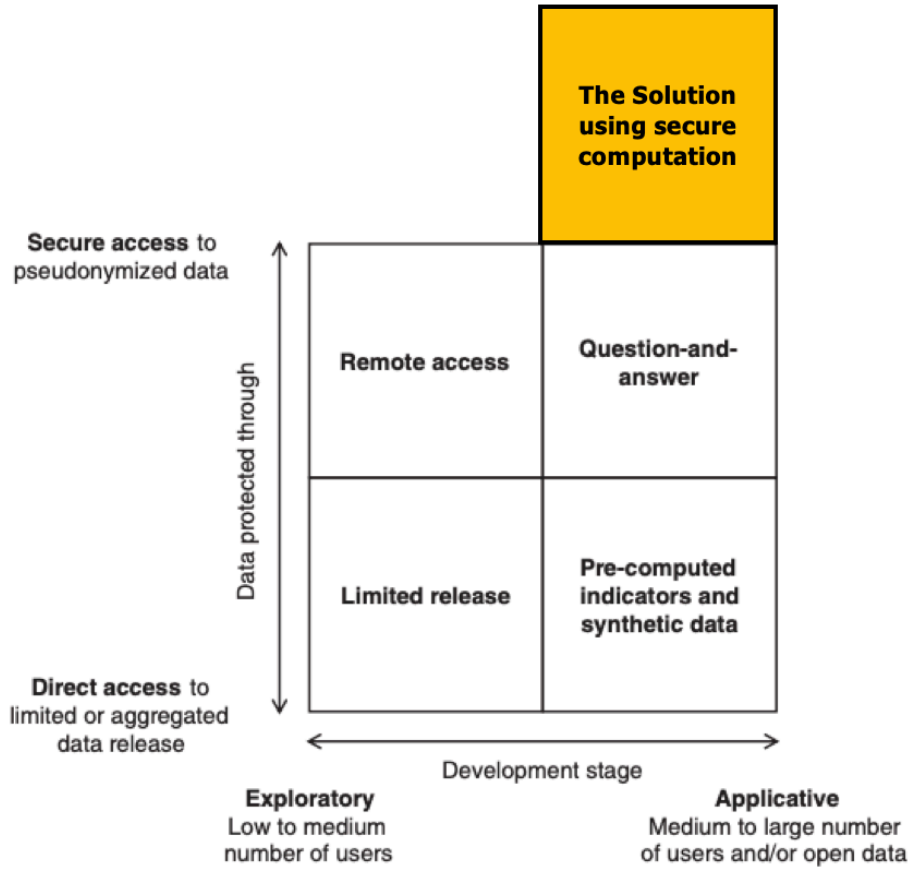
- validation of the code being used,
- a strict control of the aggregation mechanisms used for each analysis code,
- possibility to carefully add noise.

In addition, we emphasize the secure computing approach as a differentiating factor applied in the Solution. While in other models, data is being decrypted to enable computations and thus made visible to relevant stakeholders, the Solution prevents visibility of data during computations even to privileged users, thanks to the Trusted Execution Environment (TEE) component (Solution enclaves) involved in its architecture, which provides a protected memory area with confidentiality and integrity guarantees. These guarantees hold even if privileged malware is present in the system, meaning that each enclave is protected even from the operating system that is running the enclave. This way, the pseudonymised mobile location data is added an additional layer of protection which effectively secures it from being manipulated during computations, even by relevant stakeholders directly involved in the Sample Use Case. No mobile location data is disclosed to the NSI or any third parties.

As a result, we propose the Solution as a fifth model in addition to the earlier four among the matrix of the privacy-conscientious use of mobile phone data (see Figure 14 below):

Figure 14 – The Solution using secure computation as the fifth model

[342] *Ibid.*, p 4.



2) Cost of identification

Due to the application of periodically changing pseudonymization of the MNO-VAD input data in combination with other technical data protection measures (e.g. auditing of the code and enforcing that only audited code can process data, TEE for processing, and SDC for output privacy control) we argue that the cost of identification is similar to any published statistical output of NSI-s. In other words, we evaluate the risk of identification to be appropriate for producing and publishing official statistics in EU.

3) Risk of technical failures

The Sample DPIA includes technical risks analysis and also privacy risks analysis (see Section 2.5) that stem from the residual risks of the former. It concludes that compared with the situation without the Solution, only a low or very low risks are added. The fact that a few risks are added is expected as the Solution also introduces new components to provide additional features compared to the pre-existing environment.

b) Contextual elements

Due to the application of periodically changing pseudonymization of the MNO-VAD input data in combination with other technical data protection measures (e.g. auditing of the code and enforcing that only audited code can process data, TEE for

processing, and SDC for output privacy control) we argue that the risks of singling out, linkability and inference is similar to any published statistical output of NSI-s. In other words, we evaluate the risk of identification low enough in order to be appropriate for producing and publishing official statistics in EU.

ii. Legal/organisational robustness

a) Objective aspects

1) The way the processing is structured

As discussed above, the WP29 has considered it vital for the purposes of achieving isolation to ensure an adequate governance of the rights and roles for accessing personal data, which is reviewed on a regular basis. There is a need to explain how the rights and roles concept of the specific system in question ensures confidentiality of the data processed, as emphasized also in the DPIA Methodology.[343] The rights and roles concept depends on the structure of the processing, thus further clarification is needed to explain how the structure of the processing supports the rights and roles concept. This task undertaken in what follows below.

MNO-ND collects raw mobile location data for the purposes of delivering telecommunications services to Subscribers. After processing the raw mobile location data and pseudonymising the results, the MNO-ND sends the pseudonymised mobile location data to the MNO-VAD for the purposes of providing value-added services. By adopting the Solution, the NSI is also given an opportunity to process the pseudonymised mobile location data but do it in a secure environment at the MNO, kept separately from the regular business operations related to the provision of telecommunications services and value added services by the MNO.

Raw mobile location data is kept at the MNO-ND premises in non-aggregated form and relates to specific Subscribers. During pseudonymisation, the Subscribers are designated with a code. More specifically, codes are assigned to IMSIs, which indirectly relate to specific Subscribers who are using the relevant mobile devices tied to the IMSIs.[344] The MNO keeps separately the pseudonymisation key to these codes (the list associating the codes with the IMSIs). That key can be considered to be „reasonably likely to be used“ by the MNO-ND to identify Subscribers. However, since the MNO-ND also has the raw mobile location data, it can identify Subscribers through other means, irrespective of the key. Pseudonymisation is an additional protection measure to safeguard mobile location data from MNO-VAD and potential third parties. Depending on other technical and organisational measures applied at the MNO, MNO-VAD may also be considered to have „reasonably likely to be used“ means to identify Subscribers. Therefore, the set of Subscriber-related information of mobile location data can be considered as personal data subject to data protection rules by the MNO.[345]

---

[343] *Op. cit.*, F. Bieker et al. A Process for Data Protection Impact Assessment. 2016, p 32.
[344] See: Solution Architecture.
[345] *Op. cit.*, Article 29 Data Protection Working Party. Opinion 4/2007 on the concept of personal data, 2007, p 20.

The pseudonymisation key is generated and stored in an enclave on the Sharemind HI server installed at the MNO-VAD premises. It is changed periodically after each 24h period – in order to assign codes to IMSIs, the MNO-ND has to request for a new pseudonymisation key from the Solution every 24h. In response to the request, the Solution delivers a new key in encrypted from – only the MNO-ND has the means to decrypt it and is required to delete the decrypted key as soon as the pseudonymisation process is complete.

The NSI does not have access to the pseudonymised mobile location data nor to the pseudonymisation key. It can only prepare the Solution for deployment, activate it (enable the pseudonymisation process), initiate the statistical analysis process and receive the final reports. There will also be intermediate updates on the progress of the statistical analysis, but these are of purely technical nature. The Solution is designed to make sure that the NSI receives no identifiable output, even if it processes a partly aggregated version of the pseudonymised mobile location data that are initially produced by MNO-ND for MNO-VAD for the purposes of providing value-added services. WP29 has previously held that processing the same set of coded data by different controllers does not mean that each of them is processing personal data "if re-identification is explicitly excluded and appropriate technical measures have been taken in this respect"[346].

When assessing re-identifiability in the context of clinical trials, WP29 has previously stated as follows:

> *In other areas of research or of the same project, re-identification of the data subject may have been excluded in the design of protocols and procedure* [emphasis added] […]. *For technical or other reasons, there may still be a way to find out to what persons correspond what clinical data, but the identification is not supposed or expected to take place under any circumstance, and appropriate technical measures (e.g. cryptographic, irreversible hashing) have been put in place to prevent that from happening. In this case, even if identification of certain data subjects may take place despite all those protocols and measures (due to unforeseeable circumstances such as accidental matching of qualities of the data subject that reveal his/her identity), the information processed by the original controller may not be considered to relate to identified or identifiable individuals* [emphasis added] *taking account of* all the means likely reasonably to be used by the controller or by any other person [emphasis added in italics in original text]. *Its processing may thus not be subject to the provisions of the* [DPD]. *A different matter is that for the new controller who has effectively gained access to the identifiable information, it will undoubtedly be considered to be "personal data".*[347]

Since the NSI has no means „reasonably likely to be used" to identify the Subscribers and no third parties can receive any outputs from the Solution (with the only exception of MNO-VAD, which will be provided with the final reports to further

[346] *Op. cit.*, Article 29 Data Protection Working Party. Opinion 4/2007 on the concept of personal data, 2007, p 20.
[347] *Ibid.*

assure that the concerned Subscribers cannot be identified from these), the specific scheme provided by the Solution ensures that re-identification is explicitly excluded.

2) Governance of the rights and roles

The framework for defining the rights and roles concept in the Solution is based on two main sources of requirements:

1. **Legal norms**
   a. data protection law
      i. NSI as the controller of the data (see Section 8.4 below)
      ii. MNO as the processor of the data (see Section 8.4 below)
      iii. Subscriber as the data subject.
      iv. DPA as the data protection supervisor and regulator.
   b. statistics law
      i. MNO as the provider of source data
      ii. NSI as the authority with the public task to develop and produce official statistics
   c. electronic communications law
      i. MNO as the collector of the data
      ii. NSI as the secondary user of the data

2. **Technical roles**
   a. deployment planning process:
      i. MNO (MNO-VAD) as the host of the server with an Intel SGX enabled processor;
   b. Sharemind HI platform design:
      i. MNO as input provider
      ii. NSI as output consumer
      iii. Intel, Inc. as the attestation service provider.
      iv. DPA as the auditor.

Privacy-by-design and privacy-by-default principles have been implemented throughout the development of the Solution. Implementation of roles with excessive privileges has been consistently avoided. Administrators and users are given access rights to information in accordance with the least privilege principle. As a result, he key roles applied in the Solution are as follows[348]:

1) "**Input Provider**" – a task-specific role, which can upload inputs to the Solution. In the Sample Use Case, both the NSI and the MNO can provide inputs to the Solution in different processes.
2) "**Output Consumer**" – a task-specific role, which can download outputs from the Solution. In the Sample Use Case, both the NSI and the MNO (for quality control purposes) can download outputs from the Solution in different processes.
3) "**Runner**" – a task-specific role, which can start computations in the Solution. In the Sample Use Case, the MNO-ND and MNO-VAD carry out the role of the Runner in different processes.

[348] See: Solution Analysis. Sections 3.3.3., 4.3. and 4.6.

4) "**Host**" – hosts the Solution. In the Sample Use Case, the Host role is carried out by the MNO-VAD.
5) "**Coordinator**" – responsible for coordinating any setup / deployment related activities for stakeholders involved in the Solution. In the Sample Use Case, the Coordinator role is carried out by the NSI.
6) "**Developer**" – responsible for developing or ordering the development of the new/updated (statistical) analysis application. In the Sample Use Case, the Developer role is carried out by the NSI.
7) "**Auditor**" – validates critical code components *ex-ante*, before deployment, including that Solution uses the approved code and the implemented algorithm complies with privacy requirements. The Enforcers rely on this validation to approve the contents of the analysis. The Auditor role has also access to the system audit logs for *ex-post* analysis. In the Sample Use Case, the Auditor role is carried out by the NSI, the MNO-VAD (e.g. internal audit unit) and an external auditor.
8) "**Enforcer**" – required to provide approval on the contents of the analysis before the data collection or the analysis can take place. Any Enforcer can refuse new analytics to be executed if they have doubt about the kind of analysis, roles setup, or deployment that do not / no longer meet the security or privacy requirements. In the Sample Use Case, the Enforcer role is carried out by the NSI, the MNO-ND and MNO-VAD. Optionally, an external auditor could also act as an Enforcer.
9) "**Attestation Service Provider**" – proves that the expected Solution/deployment (enclave) was created on a remote machine (Host) using Intel SGX technology with the latest security patches. Before secret data is uploaded, by using remote attestation, an application can verify that a server is running trusted software in the trusted hardware. In the Sample Use Case, the Attestation Service Provider role is carried out by Cybernetica AS in the proof-of-concept stage and by Intel, Inc. in the next stages.

3) Management of shared resources

Not applicable because physical resources are not shared between different customers.

4) Risk of organizational dysfunctions (e.g. breaches of confidentiality duties)

Not applicable, presuming the mobile location data is not disclosed to the NSI and the NSI secret inputs do not contain personal data.

In case of realisation of some of the residual risks concerning an internal attacker (see the Evaluation Report), the logging mechanisms built in the Solution act as additional counter-measures against potential breaches of confidentiality duties. Legal measures and statistical quality frameworks will also act as a deterrent – statisticians are subjected to a specific duty of professional secrecy, MNO employees are subjected to obligation to keep all communication data secret. Once a breach is discovered, it will have serious implications to the relevant individuals.

5) Cost of identification

Due to the application of periodically changing pseudonymization of the MNO-VAD input data in combination with other technical data protection measures (e.g. auditing of the code and enforcing that only audited code can process data, TEE for processing, and SDC for output privacy control) we argue that the cost of identification is similar to any published statistical output of NSI-s. In other words, we evaluate the risk of identification to be appropriate for producing and publishing official statistics in EU.

b) Contextual elements

1) Content, purpose and result of processing

In the context of the Sample Use Case, the Solution releases three reports, which are aimed at producing official statistics regarding:

- **Fingerprint Report**: the pattern of population distribution depending on the time of day,
- **Population Density Report**: the typical destination zones with the level of accuracy of 1km^2,
- **FUF Report**: the approximate commuting zones of a city.

In tandem, these reports describe changes in patterns of human mobility across short distances and short periods of time. The Fingerprint Report and Population Density Report cover the whole territory of the relevant Member State in the Sample Use Case, whereas the FUF Report only covers the territories of urban areas.

In the context of the Sample Use Case, the mobile location data concerns a number of particular Subscribers in the national territory of a Member State. However, none of the stakeholders have access to or visibility of the pseudonymized mobile location data in the Solution, whether during processing, in transit or at rest. No mobile location data is disclosed to NSI or other stakeholders involved in using the Solution. This is achieved thanks to the hybrid techniques combining the following technical, legal and organisational protection measures:

1) **TEE (the Solution enclaves)** – the Solution prevents visibility of data during computations even to privileged users, thanks to the Trusted Execution Environment (TEE) component (Solution enclaves) involved in its architecture, which provides a protected memory area with confidentiality and integrity guarantees. These guarantees hold even if privileged malware is present in the system, meaning that each enclave is protected even from the operating system that is running the enclave. This way, the pseudonymised mobile location data is added an additional layer of protection which effectively secures it from being manipulated during computations, even by relevant stakeholders directly involved in the Sample Use Case. No mobile location data is disclosed to the NSI or any third parties.
2) **no transfer of personal data** – the Solution is designed in a way that the mobile location data never leaves the MNO. The NSI can process the pseudonymised mobile location data in a secure environment at the MNO,

separately from the MNO's regular business operations related to the provision of telecommunications services and value added services. The MNO keeps separately the raw (at the MNO-ND) and pseudonymised mobile location data (at the MNO-VAD), as well as the pseudonymisation keys which enable identification of individual mobile devices of Subscribers (at the MNO-ND). Moreover, the identities corresponding to the pseudonyms are kept separately from the Solution – they are stored at the MNO-ND, whereas the Solution along with the protected pseudonymization keys is hosted by the MNO-VAD. The NSI does not have access to the raw or pseudonymised mobile location data nor to the pseudonymisation keys – it can only prepare the Solution for deployment, activate it (enable the pseudonymisation process), initiate the statistical analysis process and receive the final reports.

3) **pre-approved computations** – the analysis code, including any automated controls (e.g. SDCs) are pre-approved by several independent stakeholders (NSI, MNO, potentially the local DPA or third-party auditor) and validated by an independent attestation service provider. The approval and validation procedure are technically enforced by the Solution set-up and remote attestation activities.
4) **change-and-forget pseudonymization –** raw mobile location data is pseudonymised by the MNO-ND at its premises using the periodic pseudonymisation key generated in the Solution enclave and shared with MNO-ND in encrypted form. The pseudonymisation key is changed after each 24h period. Only the MNO-ND has the means to decrypt it as per the change-and-forget method and is required to delete the decrypted key as soon as the pseudonymisation process is complete.
5) **temporal summarisation** – adding technical difficulties to attempts at re-identifying individuals and limiting the amount of information that could be uncovered if the data were to be re-identified (in Modules A, B and C).
6) **aggregation** – adding technical difficulties to attempts at re-identifying individuals and limiting the amount of information that could be uncovered if the data were to be re-identified (in Module D).
7) **governance of rights and roles** – the applied rights and roles concept limits the necessary rights and accesses in accordance with the least privilege principle.
8) **limited storage periods** – all data elements, whether in encrypted or decrypted form, are stored only until they are necessary to finalise the relevant computations. Only the 24h pseudonymisation keys (D2.1) are kept for the whole period of analysis (in the Sample Use Case deleted after 1 year) and have to be deleted manually thereafter.[349]
9) **encryption** – inputs and outputs are stored in the Trusted Execution Environment (TEE) in encrypted form, so that the data can be decrypted only inside a Solution enclave or by authorised clients.
10) **pre-approved outputs** – the output results are pre-approved by several independent stakeholders (NSI, MNO, potentially the local DPA or third-party auditor).
11) **SDC** – additional output privacy controls to assure confidentiality in accordance with applicable statistics laws.

[349] For more details on storage limitations, see 8.2.4.ii.b)4) below.

12) **auditability** – auditor(s) (e.g. data protection authorities or internal audit divisions of MNO and NSI) can be involved in the Solution development and setup to verify the correctness of the Sample Use Case Application both *ex-ante* and *ex-post*. An auditor has access to the Sample Use Case Application source code and verifies *ex-ante* the fulfilment of privacy requirements (including e.g. the non-personal nature of the final output). The auditor has also access to the system audit logs to verify *ex-post* that the data processing with the Solution was in conformity with applicable law and agreements between relevant stakeholders and that the Solution had not been tampered with (identification of potential attacks against the Solution).

These cumulative techniques enforce the data minimization principle at its maximum level, resulting in the pseudonymized mobile location data being practically anonymous throughout the processing in the Solution. As a result, the content of the mobile location data is effectively hidden from all stakeholders and information about a particular person may not be learned by any stakeholder or third party in the Solution. The the specific combination of the applied protection measures differs from one stage of processing to another, as illustrated in the Table 3 below.

Table 3 - Protection measures applied in the Sample Use Case (Meta Level)

Applied protection measures	Module A*	Module B	Module C	Module D
1) TEE (Solution enclaves)	N/A	+	+	+
2) no transfer of personal data	TBS	+	+	+
3) pre-approved computations	N/A	+	+	+
4) change-and-forget pseudonymization	+	+	N/A	N/A
5) temporal summarisation	+	+	+	N/A
6) aggregation	N/A	N/A	N/A	+
7) governance of rights and roles	TBS	+	+	+
8) limited storage periods	TBS	+	+	+
9) encryption	TBS	+	+	+
10) pre-approved outputs	TBS	N/A	N/A	+
11) SDC	N/A	N/A	N/A	+
12) auditability	N/A	+	+	+

* Module A is not part of the Solution and is thus highlighted.

Legend:
"+" – measure is applied
"N/A" – non-applicable
"TBS" – to be specified (not covered within the scope of the Sample DPIA)

A typical statistical analysis process may require collection and analysis of individualised personal information to produce the relevant statistics. Usually, traditional anonymisation techniques or other protection measures are then applied to assure statistical confidentiality at the output level. By implementing the Solution, is is possible to introduce appropriate protection measures both at the input and output level as well as during processing, in order to assure privacy at its maximum:

1) input privacy – concerns Modules A (not part of the Solution), B and C;
2) privacy during processing – concerns Modules B, C and D;
3) output privacy – concerns Module D.

With regard to the purpose of the processing, the Sample Use Case by means of the Solution is not intended to evaluate, treat in a certain way or influence the status or behaviour of a Subscriber or other individuals. Neither is it designed or used to have an impact on the rights and interests or Subscribers or other individuals. This is assured by the purpose of the processing, which is to produce official statistics.

In light of the purpose of the processing, the WP29 has considered in its Opinion on the Concept of Personal Data the following:

> *Where identification of the data subject is not included in the purpose of the processing, the technical measures to prevent identification have a very important role to play. Putting in place the appropriate state-of-the-art technical and organizational measures to protect the data against identification may make the difference to consider that the persons are not identifiable, taking account of* all the means likely reasonably to be used by the controller or by any other person [emphasis provided in italic in original text] *to identify the individuals. In this case, the implementation of those measures are not the* consequence [emphasis provided in italic in original text, underline added] *of a legal obligation arising from Article 17 of the* [DPD] *(which only applies if the information is personal data in the first place), but rather a* condition [emphasis provided in italic in original text, underline added] *for the information precisely not to be considered to be personal data and its processing not to be subject to the* [DPD]*.*[350]

The above statement of WP29 is relevant for the purposes of the Sample DPIA, if the further processing pseudonymous mobile location data by means of the Solution for producing official statistics is considered as “making anonymous” under ePD Art 9(1) and this is confirmed by relevant data protection authorities and courts. In such case, the Solution functions as a condition for the further processing. In other cases, the Solution can be applied as a set of appropriate safeguards for statistical purposes under GDPR Art 89(1).

2) Expected advantage to the controllers vs interests of individuals

The novel kind of information provided in the reports released by the Solution in the Sample Use Case helps the NSIs better fulfil their public task of providing relevant and accurate statistics for a range of public policy goals. It can also be valuable for

[350] *Op. cit.*, Article 29 Data Protection Working Party. Opinion 4/2007 on the concept of personal data, 2007, p 17.

organisations in the private sector. For example, by having more accurate information about the commuting corridors, the local administrations can plan better public transport lines and companies can optimise their transportation schedules and costs.

Considering that the purpose of the processing is to produce official statistics, not provide any individual effects or otherwise create individual impacts, the interests of individuals are expected to remain untouched. At the same time, individuals as members of the public and residents of relevant cities and local administrations will share benefits deriving from the public authorities making more reliable policy decisions based on the reports released by the Solution in the Sample Use Case.

Based on the above considerations, the expected advantage of the controllers can be considered to be in balance with the interests of individuals. However, the balance will need to be re-evaluated for each statistical analysis use case developed for real-world scenarios in the future, based on its own risks and merits.

3) Rarity of phenomenon (including population density, nature and volume of data)

It is anticipated that further processing pseudonymous mobile location data by means of the Solution in the Sample Use Case can reflect rarities of different phenomena (e.g. small number of residents in remote areas of the country or secluded areas in the city). To counter the risk of identification in these sensitive cases, the following protection measures have been implemented:

- periodical pseudonymisation,
- temporal summarisation,
- aggregation,
- SDC.

Provided that the SDC are customised to the peculiarities of the relevant Member State in the Sample Use Case, these measures can be considered to be satisfactory to counter the risk of identification in case of certain rare phenomena.

4) Data storage

In the context of the Sample Use Case, the pseudonymised mobile location data are collected over a period of one year. After collection, the pseudonymised mobile location data are stored at rest in encrypted form, capable of being decrypted only within the enclave of the Solution during computations for the Sample Use Case.

The following tables in this section summarise the different types of mobile location data that are received by the relevant stakeholders (NSI and MNO) for further processing in the context of the Sample Use Case. The types of personal data processed have been divided into two tables:

1) Table 4 represents the types of personal data relevant in the Primary Processing phase,
2) Table 5 represents the types of personal data relevant in the Secondary Processing phase.

All data elements, whether in encrypted or decrypted form, are stored only until they are necessary to finalise the relevant computations (see Table 4 and Table 5 – Data storage in Secondary Processing below). Only the 24h pseudonymisation keys (D2.1) are kept for the whole period of analysis (in the Sample Use Case deleted after 1 year) and have to be deleted manually thereafter.

Table 4 – Data storage in Primary Processing

Data types	Recipients	Storage duration
Raw mobile location data (D2.4)	MNO-ND	D2.4 is deleted by the MNO-ND at their discretion. Defined by: 1) applicable electronic communications and data protection laws and regulations. 2) MNO internal business processes (MNO-ND).
24h pseudonymisation keys (D2.1, D2.2, D2.3)	Trusted Execution Environment (generates, stores, processes and protects the keys)	D2.1 is stored until the period of analysis expires and all data of the Trusted Execution Environment is deleted manually (in the Sample Use Case after 1 year). D2.2 is an encrypted copy of one element from D2.1. D2.2 is sent to the MNO-ND, who decrypts it and obtains D2.3. Defined by: 1) applicable statistics and data protection laws and regulations, 2) contractual arrangements between the NSI and the MNO, 3) settings of the MNO systems (e.g. back-up frequency etc), 4) updates to the Solution.
	MNO-ND	D2.2 and D2.3 are deleted by the MNO-ND right after use in accordance with the change-and-forget method. Defined by: 1) applicable electronic communications and data protection laws and regulations, 2) MNO internal business processes (MNO-ND).
Pseudonymised mobile location data (D2.5, D2.6)	MNO-ND	D2.5 is deleted by the MNO-ND at their discretion. Defined by: 1) applicable electronic communications and data protection laws and regulations.

		2) MNO internal business processes (MNO-ND).
	MNO-VAD	D2.6 is deleted by the MNO-VAD at their discretion. Defined by: 1) applicable electronic communications and data protection laws and regulations. 2) MNO internal business processes (MNO-VAD).

Table 5 – Data storage in Secondary Processing

Data types	Recipients	Storage duration
Pseudonymised mobile location data (D2.6)	MNO-VAD (sends to the Module A)	D2.6 is deleted by the MNO-VAD right after relevant computations on it (P5.1 on Figure 11) have been completed in the Module A. Defined by: 1) applicable electronic communications, statistics and data protection laws and regulations, 2) contractual arrangements between the NSI and the MNO, 3) settings of the MNO systems (e.g. back-up frequency etc).
Temporally summarised pseudonymised mobile location data (D5.1)	MNO-VAD (generates and stores the data in the Module A)	D5.1 is deleted by the MNO-VAD right after importing it from the Module A to the Solution. Defined by: 1) applicable statistics and data protection laws and regulations, 2) contractual arrangements between the NSI and the MNO, 3) settings of the MNO systems (e.g. back-up frequency etc).
	Trusted Execution Environment (processes and protects the data received from the Module A)	D5.1 is deleted right after it has been reverse pseudonymised in the Solution's enclave and relevant computations on it (P5.3 on Figure 11) have been completed. Defined by: 1) applicable statistics and data protection laws and regulations, 2) contractual arrangements between the NSI and the MNO,

		3) settings of the MNO systems (e.g. back-up frequency etc), 4) updates to the Solution.
24h pseudonymisation keys (D2.1)	Trusted Execution Environment (generates, stores, processes and protects the keys)	D2.1 is deleted manually after the period of analysis expires and all data of the Solution enclave is deleted (in the Sample Use Case after 1 year). Defined by: 1) applicable statistics and data protection laws and regulations, 2) contractual arrangements between the NSI and the MNO, 3) settings of the MNO systems (e.g. back-up frequency etc), 4) updates to the Solution.
Temporally summarised reverse pseudonymised mobile location data (D5.2)	Trusted Execution Environment (processes and protects the data)	D5.2 is deleted right after the relevant computations on it (P5.3 on Figure 11) have been completed. Defined by: 1) applicable statistics and data protection laws and regulations, 2) contractual arrangements between the NSI and the MNO, 3) settings of the MNO systems (e.g. back-up frequency etc), 4) updates to the Solution.

The period of one year chosen for the proof-of-concept setting of the Sample Use Case is long enough to allow TEE-based privacy-enhancing technologies to mature and it can be foreseen that there will be more of similar offerings available in a year. However, this does not mean the frontline of the state-of-the-art in this field will necessarily move forward considerably over the period of one year. For this reason, the Solution can be foreseen to maintain its robustness and identification of Subscribers is not anticipated to be possible during the lifetime of the mobile location data in the Solution.[351]

3) Possibility to combine the personal data elements held by different persons

According to the ECJ judgment in the Breyer case, there is no possibility to combine the personal data elements held by different persons and thus identify a data subject if the identification of the data subject was:

1. prohibited by law or

[351] *Op. cit.*, Article 29 Data Protection Working Party. Opinion 4/2007 on the concept of personal data, 2007, p 15.

2. practically impossible on account of the fact that it requires a disproportionate effort in terms of time, cost and man-power, so that the risk of identification appears in reality to be insignificant.

Mobile location data is considered to be very sensitive. A strict legal regime has been built under the EU law to protect it. In principle, besides the Subscribers themselves, only MNOs are allowed to process mobile location data for providing electronic communications services, as outlined in ePD Art 9. More specifically, ePD Art 9(1) prohibits MNO from sharing mobile location data with any third parties or otherwise further processing it without anonymising it first. This means that re-use of mobile location data is in general prohibited, save for the specific exceptions allowed in ePD:

1. ePD Art 9(1) second alternative allows re-use of mobile location data for enabling value-added-services to Subscribers,

2. ePD Art 10(2) provides exceptions for organisations dealing with emergency calls and recognised as such by a Member State, including law enforcement agencies, ambulance services and fire brigades, for the purpose of responding to such calls,

3. ePD Art 15 allows Member States to adopt legislative measures providing for the retention of data, including mobile location data, for a limited period justified on the grounds of, *inter alia*, safeguarding national security (i.e. State security), defence, public security, and the prevention, investigation, detection and prosecution of criminal offences or of unauthorised use of the electronic communication system. However, such restrictions must constitute a necessary, appropriate and proportionate measure within a democratic society and be in accordance with the general principles of EU law, including the general principles and fundamental rights now guaranteed by the Charter. This exception has been interpreted by the ECJ to apply in three areas:

    a. **Civil proceedings**. The ECJ has confirmed that ePD Art 15(1), in conjunction with relevant clauses of Directive 2004/48/EC of the European Parliament and of the Council of 29 April 2004 on the enforcement of intellectual property rights, does not preclude Member States from laying down an obligation for MNOs to disclose to private third parties personal data relating to Internet traffic to enable them to initiate civil proceedings for copyright infringements. However, other acts of EU law require Member States to ensure that they rely on an interpretation of those directives which allows a fair balance to be struck between the various fundamental rights protected by the EU legal order.[352]

[352] European Court of Justice. Order of the Court (Eighth Chamber) of 19 February 2009. LSG-Gesellschaft zur Wahrnehmung von Leistungsschutzrechten GmbH v Tele2 Telecommunication GmbH. Case C-557/07, European Court Reports 2009 I-01227. – Internet: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A62007CO0557 (11.05.2021).

b. **Criminal proceedings**. The ECJ has concluded that "[t]he protection of the confidentiality of electronic communications and related traffic data, guaranteed in Article 5(1) of [ePD], applies to the measures taken by all persons other than users, whether private persons or bodies or State bodies. As confirmed in recital 21 of that directive, the aim of the directive is to prevent unauthorised access to communications, including 'any data related to such communications', in order to protect the confidentiality of electronic communications."[353] Furthermore, "[t]he scope of Article 5, Article 6 and Article 9(1) of [ePD], which seek to ensure the confidentiality of communications and related data, and to minimise the risks of misuse, must moreover be assessed in the light of recital 30 of that directive, which states: 'Systems for the provision of electronic communications networks and services should be designed to limit the amount of personal data necessary to a strict minimum'."[354] At the same time, ePD Art 15(1) "enables the Member States to introduce exceptions to the obligation of principle, laid down in Article 5(1) of that directive, to ensure the confidentiality of personal data, and to the corresponding obligations, referred to in Articles 6 and 9 of that directive."[355] However, ePD Art 15(1) must be interpreted strictly and it cannot permit the exception to become the rule.[356] According to the ECJ, Member States are allowed to adopt "legislation permitting, as a preventive measure, the targeted retention of traffic and location data, for the purpose of fighting serious crime, provided that the retention of data is limited, with respect to the categories of data to be retained, the means of communication affected, the persons concerned and the retention period adopted, to what is strictly necessary."[357]

c. **National security**. Member States are obliged to comply with EU law also in situations where a national measure has been taken for the purpose of protecting national security.[358] Even though ePD Art 15(1) allows exemptions from data protection law for the purposes of national security, all personal data processing operations carried out by MNOs fall within the scope of ePD, including processing operations resulting from obligations imposed on MNOs by the public authorities.[359] In conclusion, national legislation enabling a State authority to require MNOs to forward traffic data and location data to the security and intelligence agencies for the purpose of safeguarding national security

[353] *Op. cit.*, European Court of Justice. Judgment of the Court (Grand Chamber) of 21 December 2016. Tele2 Sverige AB v Post- och telestyrelsen and Secretary of State for the Home Department v Tom Watson and Others. Joined Cases C-203/15 and C-698/15, sec 77.
[354] *Ibid.*, sec 87.
[355] *Ibid.*, sec 88.
[356] *Ibid.*, sec 89.
[357] *Ibid.*, sec 108.
[358] European Court of Justice. Judgment of the Court (Grand Chamber) of 6 October 2020. Privacy International v Secretary of State for Foreign and Commonwealth Affairs and Others. Case C-623/17, sec 44. – Internet: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A62017CJ0623 (13.05.2021).
[359] *Ibid.*, sec 46.

> falls within the scope of ePD.[360] The prohibition on the interception of communications and data relating thereto "encompasses any instance of providers of electronic communications services making traffic data and location data available to public authorities, such as the security and intelligence agencies, as well as the retention of that data by those authorities, regardless of how that data is subsequently used."[361] However, ePD Art 15(1) provides for exceptions to this rule. "That being said, the option to derogate from the rights and obligations laid down in Articles 5, 6 and 9 of [ePD] cannot permit the exception […] to become the rule".[362] Such exceptions must fulfil the criteria set out in Art 52(1) of the Charter: "provided that those limitations are provided for by law, that they respect the essence of those rights and that, in compliance with the principle of proportionality, they are necessary and genuinely meet objectives of general interest recognised by the Union or the need to protect the rights and freedoms of others".[363] More specifically, "the legal basis which permits the interference with those rights must itself define the scope of the limitation on the exercise of the right concerned".[364] According to ECJ, "in order to satisfy the requirement of proportionality, according to which derogations from and limitations on the protection of personal data must apply only in so far as is strictly necessary, national legislation entailing interference with the fundamental rights enshrined in Articles 7 and 8 of the Charter must meet the requirements stemming from the case-law cited in paragraphs 65, 67 and 68 [of the judgment]."[365]

Based on the exceptions addressed in ECJ case law, there is a chance that the MNO might be requested for mobile location data of a Subscriber or a group of Subscribers under national law of the relevant Member State. Even so, for the purposes of this document, it is not realistic that such request comes from any of the stakeholders involved in using the Solution. Subscribers do not have an active role in the Sample Use Case, which could give rise to their mobile location data being retained or processed on the grounds referred to in ePD Art 15(1). Mobile location data relevant for these purposes can be obtained through other means, but not through the Solution.

In the Breyer case, the ECJ considered the possibility to obtain, with the assistance of other persons, the missing elements of personal data when initiating criminal proceedings in the event of cyber attacks, and qualified it as "the means which may likely reasonably be used to identify the data subject".[366] In the context of the Sample Use Case, this legal route does not apply for the following reasons:

---

[360] *Ibid.*, sec 19.
[361] *Ibid.*, sec 56.
[362] *Ibid.*, sec 59.
[363] *Ibid.*, sec 64.
[364] *Ibid.*, sec 65.
[365] *Ibid.*, sec 76.
[366] *Op. cit.*, European Court of Justice. Judgment of the Court (Second Chamber), 19 October 2016, C-582/14, Patrick Breyer v Bundesrepublic Deutschland, sec 47-48.

1. Unlike in the Breyer case, there is no need for the stakeholders to identify any individual data subjects. On the contrary, due to the obligation to assure statistical confidentiality and the accompanying duties down to the level of an individual statistical analyst, the NSI is obliged to refrain from identifying individuals and, where this is inevitable, keep the identifying data confidential.

2. The result of processing pseudonymized mobile location data for the purposes of producing official statistics by means of the Solution is aggregated data and, thanks to applying the hybrid techniques (see Section 8.2.4 above), most probably qualifies as anonymous data. The NSI is not allowed to use any results of statistical analysis, whether anonymous or not, in support of measures or decisions regarding any particular natural person. Therefore, it is prevented from requesting additional information from MNO, which would allow it to reverse the pseudonyms and re-identify relevant Subscribers.

3. The Solution is designed to have no bearing on the identifiability of Subscribers. The MNO can identify Subscribers based on the raw mobile location data and client relationship management data collected as part of regular service offering to its Subscribers, irrespective of the Solution. Other stakeholders have no access rights, technical means or interest in obtaining personal data about Subscribers by means of the Solution.

4. The pseudonymous nature of the pseudonymized mobile location data prevents it from identifying the underlying data subject without additional information. Such additional information – the pseudonymization keys and the relevant identities corresponding to the pseudonyms – are kept in the premises of the MNO. Moreover, the identities corresponding to the pseudonyms are kept separately from the Solution – they are stored at the MNO-ND, whereas the Solution along with the protected pseudonymization keys is hosted by the MNO-VAD. The Solution is designed in a way that the MNO-ND cannot share the identity and pseudonymisation key information with the MNO-VAD and/or the NSI, which would allow them to re-identify Subscribers.

5. Under current laws (*de lege lata*), the NSIs and other stakeholders do not have a legal route to obtaining mobile location data directly from the Solution because ePD Art 9(1) prevents them from doing so. Any relevant national laws at Member State level must respect the limitation of ePD Art 9(1) because it is *lex specialis* in relation to GDPR.

8.2.5. Interim conclusion

Thanks to introducing a new state-of-the-art, the Solution provides a level of robustness which is superior to earlier models of privacy-conscientious processing of mobile location data in several respects. The further processing of pseudonymous mobile location data by means of the Solution for producing official statistics can be

interpreted as “made anonymous” in terms of ePD Art 9(1) due to the application of a complex hybrid of technical, organisational and legal measures (hybrid techniques).

As a result of applying the hybrid techniques in the Solution, the pseudonymised mobile location data are effectively “made anonymous” during a process which:
1) starts as of the moment of encrypting the pseudonymised mobile location data data for the Trusted Execution Environment (P5.2 on Figure 11) and
2) ends when the pseudonymous mobile location data has been temporally summarised and aggregated and the the relevant computations on it have been completed within the Trusted Execution Environment (P5.3 on Figure 11).

The objective of implementing the Solution is to ensure that the risk of reidentification of the data subjects is minimal. Identification of Subscribers by the stakeholders as well as any third parties by means of the Solution is prohibited by law and practically impossible, so that the risk of identification appears in reality to be insignificant. All facts confirm very low or minimal risks of re-identification and relatively low impact on the Subscribers if it nevertheless happens[367]. Therefore, processing pseudonymous mobile location data by means of the Solution for producing official statistics can be qualified as “made anonymous” under ePD Art 9(1). Consequently, it fulfils not only the requirements of the lower threshold for the notion of anonymity under statistics law but also the higher threshold for the same in data protection law under applicable law and regulations as they stand at the moment of completing the Sample DPIA (see Section 7.2.3 above).

If the relevant data protection authorities and courts agree that the further processing of pseudonymous mobile location data by means of the Solution for producing official statistics is considered as “making anonymous” under ePD Art 9(1), then the Solution functions as a condition for the further processing (see Section 8.2.4.ii.b)1). This means that the condition of “making anonymous” under ePD Art 9(1) holds if and only if no personal data is extracted from the Solution nor shared with the stakeholders or third parties by means of the Solution. This includes, *inter alia*:
1) no extracting of pseudonymisation keys which could be used for reverse pseudonymising the pseudonymous mobile location data at the MNO;
2) no individual Subscribers are identifiable from the output results of the Solution.

Since the the answer to the first question is “yes”, then the second question was whether further processing pseudonymous mobile location data by means of the Solution for producing official statistics is carried out by persons under the authority of the MNO, in order to meet the additional requirement of further processing set out in ePD Art 9(3). Considering that the Solution is installed at the MNO premises and assuming only personnel authorised by the MNO have access to it, in accordance with the relevant rights and roles concept, the second question is also to be answered as a “yes”.

This means that further processing pseudonymous mobile location data by means of the Solution for producing official statistics is in compliance with ePD Art 9 and the MNO is allowed according to *de lege lata* to make this data available to the NSI by

[367] See: Evaluation Report Section 4.2.2 for the final conclusion regarding the level of residue risks.

means of the Solution, provided there is a proper legal basis for it under GDPR Art 6 (see Section 8.5 below).

This conclusion may have wide-spread practical implications for different statistical production processes in the future. A typical statistical analysis process may require collection and analysis of individualised personal information to produce the relevant statistics. Usually, traditional anonymisation techniques or other protection measures are then applied to assure statistical confidentiality. The Solution is designed to demonstrate the opposite – statistical authorities no longer need to collect, store and analyse individualised personal information inhouse in order to produce statistics in accordance with the statistical quality principles. Thus, the question of necessity becomes central when considering a legal basis on the ground of public interest/official authority – it is required under GDPR Art 6(3) that the processing must be necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller.

The ECJ has analysed the concept of necessity in the context of DPD Art 7(e), which provided that personal data may lawfully be processed 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 or in a third party to whom the data are disclosed". The case concerned the statistical function of a Central Register of Foreign Nationals (Ausländerzentralregister) ("the **AZR**").[368] On the one hand, the ECJ recalled that, EU law has not excluded the power of Member States to adopt measures enabling the national authorities to have an exact knowledge of population movements affecting their territory, referring to an earlier judgment in the case Watson and Belmann.[369] On the other hand, the ECJ concluded that "the storage and processing of personal data containing individualised personal information in a register such as the AZR for statistical purposes cannot, on any basis, be considered to be necessary within the meaning of Article 7(e) of [DPD]"[370] because "it is only anonymous information that requires to be processed"[371] in order to collect information and determine statistics relating to migratory flows in the territory of Member States with the objective of transmitting those statistics in accordance with Regulation No 862/2007.[372]

Therefore, if further processing pseudonymous mobile location data by means of the Solution for producing official statistics in the context of the Sample Use Case qualifies as anonymous data, it raises the necessity bar in terms of producing the same statistics from other information which is collected for the Sample Use Case purposes, as long as the Solution is established as a new state-of-the-art.

[368] European Court of Justice. Judgment of the Court. 16 December 2008, In Case C-524/06, Heinz Huber v Bundesrepublik Deutschland. – Internet: https://curia.europa.eu/juris/document/document.jsf?text=&docid=76077&pageIndex=0&doclang=EN&mode=lst&dir=&occ=first&part=1&cid=13821220 (03.05.2021).
[369] *Ibid.*, sec 63.
[370] *Ibid.*, sec 68.
[371] *Ibid.*, sec 65.
[372] *Ibid.*, sec 64-65.

As the final steps for completing the Sample DPIA, three assessments will be carried out for further processing pseudonymous mobile location data by means of the Solution for producing official statistics:

1. **compatibility assessment** (limited to assessing the appropriateness of the proposed safeguards) (see Section 8.3 below),
2. **controllership assessment** (see Section 8.4 below),
3. **lawfulness assessment** (see Section 8.5 below).

8.3. Compatibility assessment

Based on the discussion above, it can be concluded that further processing pseudonymous mobile location data by means of the Solution for producing official statistics can be interpreted as "making anonymous" in terms of ePD Art 9(1) (see Section 8.2.5 above). Since ePD must be construed in line with GDPR, it was also concluded that "make anonymous" under ePD Art 9(1) and "render anonymous" under GDPR Rec 26 should be treated as synonyms, i.e. they denote the same concept (see Section 8.2.1 above). For ease of reference, the concepts of "make anonymous" and "render anonymous" shall be jointly addressed as "anonymisation" hereafter. It is assumed that the relevant guidelines of data protection authorities, both in pre-GDPR era and thereafter, take the same approach and, thus, there is no need to distinguish between the different meanings of the vocabulary used to denote the same concept.

Some authors have pointed out that there is legal uncertainty regarding the lawfulness of the anonymisation process.[373] The WP29 Opinion on Anonymisation Techniques took the position that anonymisation is a type of data processing. It thus assumed that the personal data must have been collected and processed in compliance with the applicable legislation on the retention of data in an identifiable format. In this context, the processing of such personal data to achieve their anonymisation is an instance of "further processing", which must comply with the test of compatibility in accordance with the WP29 guidelines provided in its Opinion on Purpose Limitation.[374]

According to WP29, "anonymisation as an instance of further processing of personal data can be considered to be compatible with the original purposes of the processing but only on condition the anonymisation process is such as to reliably produce anonymised information in the sense described in [WP29 Opinion on Anonymisation Techniques]."[375] At the same time, there is no express presumption of compatibility foreseen in GDPR for anonymisation *per se*. Therefore, the WP29 interpretation of anonymisation as compatible further processing should be treated as a mere possibility that needs to be checked against the compatibility test under GDPR Art

[373] *Op. cit*., G. Spindler, P. Schmechel. Personal Data and Encryption in the European General Data Protection Regulation, 2016, p 173. See also: G. Spindler, A. Z. Horváth. Deliverable D3.5 Use-case specific legal aspects. Scalable Oblivious Data Analytics (SODA), p 30. – Internet: https://soda-project.eu/wp-content/uploads/2019/12/SODA-D3.5-Use-case-specific-legal-aspects.pdf (03.05.2021).
[374] *Op. cit*, Article 29 Data Protection Working Party, Opinion 05/2014 on Anonymisation Techniques, 2014, p 7.
[375] *Ibid.*

6(4). This is not the case when anonymisation is conducted for archiving purposes in the public interest, scientific research or statistical purposes, because here the presumption of compatibility applies under GDPR Art 5(1)(b). As concluded above, an assessment of appropriate and sufficient safeguards is assumed to suffice as a minimum requirement for applying the presumption of compatibility. The latter exercise is undertaken in the following section.

For the purposes of the Project, the mobile location data is processed in order to produce official statistics. This purpose triggers the application of the presumption of compatibility under GDPR Art 5(1)(b). Therefore, in the context of the Sample Use Case, anonymisation of mobile location data for statistical purposes is not considered incompatible with the initial purposes of collecting the mobile location data, if appropriate safeguards are implemented in compliance with GDPR Art 89(1). Since ePD Art 9(1) expressly requires "making anonymous" of mobile location data before it is further processed, the question is whether fulfilling this requirement also fulfils the criteria of appropriate safeguards under GDPR Art 89(1). In other words, does processing of mobile location data by means of the Solution qualify as appropriate safeguards in terms of GDPR Art 89(1), and, consequently, justify the application of the presumption of compatibility?

According to WP29 guidance from the pre-GDPR era, a generally applicable multi-factor approach should be applied in order to identify the appropriate safeguards.[376] In order for appropriate safeguards to serve as compensation for a change of purpose, e.g. in case of further processing, technical and/or organisational measures might be required to ensure functional separation but also additional steps might be required to be taken for the benefit of the data subjects, such as increased transparency, with the possibility to object or provide specific consent. "Whether the result is acceptable will depend on the compatibility assessment as a whole (i.e. including those measures and their effect on the other aspects mentioned above [referring to a) the relationship between the purposes for which the data have been collected and the purposes of further processing, b) the context in which the data have been collected and the reasonable expectations of the data subjects as to their further use, c) the nature of the data and the impact of the further processing on the data subjects])."[377] "[T]he easier the data subject can be identified, the more additional safeguards will be needed."[378] Therefore, the assessment of appropriate safeguards should be carried out in two steps:[379]

1) assessment of the possibilities and limits of effective de-identification,
2) applying additional safeguards.[380]

In what follows, we shall implement this two-step approach based on examples of its application by the WP29, which are referred to accordingly. It is assumed that a careful impact assessment is made, penetration tests are carried out, and stakeholders are consulted in the pilot project and production stages.

---

[376] *Op cit.*, Article 29 Data Protection Working Party. Opinion 03/2013 on purpose limitation, 2013, p 26, 30, 33, 56,
[377] *Ibid.*, p 26.
[378] *Ibid.*, p 32.
[379] *Ibid*.
[380] *Ibid.*

8.3.1. Possibilities and limits of effective de-identification

Mobile location data initially collected and pseudonymised for a specific purpose by the MNO (delivering telecommunications services to Subscribers) are now used for different purposes – producing official statistics. Most people would not commonly expect their data to be used in this way, which gives a strong indication that the initial and secondary purposes are incompatible. This assessment is also supported by the high sensitivity of mobile location data.[381]

However, in this case, as part of the re-use for the secondary purpose, the mobile location data is effectively anonymised by means of hybrid techniques (see Section 8.2.5 above). All facts confirm very low or minimal risks of re-identification and relatively low impact on the data subjects if it nevertheless happens. Therefore, although the two purposes are different, the applied hybrid techniques reduce any concerns regarding incompatible processing.[382]

8.3.2. Additional safeguards

Nevertheless, additional safeguards, such as full transparency about the processing are still recommended.[383] This will be achieved by fulfilling the requirements outlined in ESCoP, which may be further specified in national law of the relevant Member State:

1. **Principle 4 "Commitment to Quality"**
   Indicator 4.1: Quality policy is defined and made available to the public. An organisational structure and tools are in place to deal with quality management.

2. **Principle 5 "Statistical Confidentiality and Data Protection"**
   Indicator 5.4: Guidelines and instructions are provided to staff on the protection of statistical confidentiality throughout the statistical processes. The confidentiality policy is made known to the public.

3. **Principle 6 "Impartiality and Objectivity"**
   Indicator 6.4: Information on data sources, methods and procedures used is publicly available.

As part of the Sample DPIA process, a technical and privacy risk analysis was carried out. The conclusions of this risk analysis are documented in the Evaluation Report, which will be published as a result of the Project (see Section 8.2.5 above). According to the risk analysis, there are some residual risks, also in terms of potential impact to the fundamental rights and freedoms of data subjects. However, they are considered to be low or very low.[384]

---

[381] See: Annex 4, Example 15 – *Op cit.*, Article 29 Data Protection Working Party. Opinion 03/2013 on purpose limitation, 2013, p 66.
[382] *Ibid.*, pp 66-67.
[383] *Ibid.*, p 67.
[384] *Ibid.*

Additionally, as a rule, informed consent will be required.[385] The Subscribers' expectation of consent has also been confirmed by ECJ: "the users of electronic communications services are entitled to expect, in principle, that their communications and data relating thereto will remain anonymous and may not be recorded, unless they have agreed otherwise."[386] This means that before the start of the processing, the controller(s) will need to obtain a consent from relevant Subscribers to process their mobile location data by means of the Solution for producing official statistics.

Alternatively, if the legal basis for further processing pseudonymous mobile location data by means of the Solution for producing official statistics is grounded in EU law or national law of the relevant Member State (see Section 8.5 below) which contains exceptions to the requirement of consent along with appropriate safeguards, it may need to provide an opportunity to allow Subscribers to opt-in or opt-out of the processing along with a notification to or authorisation of the relevant data protection authority.[387] The applicability of this safeguard will require further legal analysis based on the actual statistical analysis use cases to be developed in the future.

Under GDPR, just as under DPD, it is up to each Member State to specify what safeguards may be considered as appropriate. Typically, this specification is provided in legislation, which could be precise (e.g. national census or other official statistics) or more general (most other kinds of statistics or research). "In the latter case, this leaves room for professional codes of conduct and/or further guidance released by competent data protection authorities."[388]

8.4. Controllership assessment

In September 2020, the EDPB published its new Guidelines 07/2020 on the concepts of controller and processor in the GDPR for public consultation,[389] followed by a finally adopted version after public consultation on 7 July 2021[390] ("**EDPB Controllership Guidelines**"). It incorporates interpretations of recent case law in the matter and thus provides an important guidance on how to assign controllership under the GDPR.

---

[385] *Ibid.*
[386] *Op. cit*. European Court of Justice. Judgment of the Court (Grand Chamber) of 6 October 2020. Privacy International v Secretary of State for Foreign and Commonwealth Affairs and Others. Case C-623/17, sec 57.
[387] *Op cit.*, Article 29 Data Protection Working Party. Opinion 03/2013 on purpose limitation, 2013, p 26, p 32.
[388] *Ibid.*
[389] European Data Protection Board. Guidelines 07/2020 on the concepts of controller and processor in the GDPR Version 1.0 Adopted on 02 September 2020. Adopted – version for public consultation. – Internet: https://edpb.europa.eu/our-work-tools/documents/public-consultations/2020/guidelines-072020-concepts-controller-and_en (13.05.2021).
[390] European Data Protection Board. Guidelines 07/2020 on the concepts of controller and processor in the GDPR Version 2.0 Adopted on 07 July 2021. Adopted – After public consultation. – Internet: https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-072020-concepts-controller-and-processor-gdpr_en (09.08.2021).

The EDPB Controllership Guidelines propose 5 building blocks, when analysing controllership:

1. **type of entity that can be a controller** – in principle, there is no limitation as to the type of entity that may assume the role of a controller, i.e. it can be an organisation, an individual or a group of individuals.[391]
2. **"determines"** – refers to the influence over the processing by virtue of exercise of decision-making power – a controller is a body that decides certain key elements about the processing. It is based on a factual rather than a formal analysis, providing answers to questions "why is this processing taking place?" and "who decided that the processing should take place for a particular purpose?". Further, the EDPB distinguishes between two categories of situations:[392]
    a. **control stemming from legal provisions** – here, control can be inferred from explicit legal competence. This is the case where
        i. **the controller has been specifically identified by law** (controller is designated as "the entity that has genuine ability to exercise control", e.g. national law providing that a public authority is responsible for processing personal data within the context of its duties)[393] or
        ii. **the law will establish a task or impose a duty on someone to collect and process certain data** (controller is designated by law for the realization of the purpose or public task determined by the law). If a law imposes an obligation on either public or private entities to retain or provide certain data, these entities would then normally be considered as controllers with respect to the processing that is necessary to execute this obligation.[394]
    b. **control stemming from factual influence** – in the absence of control arising from legal provisions, controllership is established on the basis of an assessment of the factual circumstances surrounding the processing. All relevant circumstances have must be taken into account. The qualification as controller or processor has to be assessed with regard to each specific data processing activity. An assessment of the contractual terms between the different parties involved can facilitate the determination, if the contract accurately reflects the reality.[395]
3. **"alone or jointly with others"** – an organisation can still be a controller even if it does not make all the decisions as to purposes and means.[396]
4. **"purposes and means"** – this refers to the object of the controller's influence, i.e. what a party should determine in order to qualify as controller. Determining the purposes and means amounts to deciding respectively the "why" and the "how" of the processing. When drawing the line between decisions that are reserved to the controller and those left to the processor, controller makes decisions on the purpose of the

[391] *Ibid.*, sec 2.1.1.
[392] *Ibid.*, Sec 2.1.2.
[393] *Ibid.*, Sec 2.1.2, 1), p 11.
[394] *Ibid.*
[395] *Ibid.*, Sec 2.1.2, 2), pp 12-13.
[396] *Ibid.*, Sec 2.1.3, p 14.

processing and decisions on the "essential means" ("which data shall be processed?", "who shall have access to them?", "whose personal data are being processed?"), whereas processor may decide on "non-essential means" (more practical aspects of implementation, such as the choice of a particular type of hard- or software or the detailed security measures). At the same time, controller must be fully informed about the means that are used so that it can take an informed decision in this regard and demonstrate the lawfulness of the processing.[397]

5. **"of the processing of personal data"** – the concept of a controller can be linked either to a single processing operation or a set of operations. The control exercised by one entity may be limited to a particular stage in the processing. There is a need to differentiate between "micro-level" and "macro-lecel" – while at "micro-level" the different processing operations of a chain may appear as disconnected, they may be considered as a "set of operations" pursuing a joint purpose using jointly defined means at "macro-level." It is not required that the controller actually has access to the data that is being processed.[398]

Analysing these building blocks in the context of deploying the Solution, it is assumed that the NSI contracts MNO as a service provider to obtain the relevant statistical analysis reports, unless the national law of the relevant Member State has imposed a duty on the MNO to process mobile location data by means of the Solution for producing official statistics. In this setting, it is clear that the NSI instructs the MNO on what type of information it is interested in and provides the algorithms to run the statistical analysis. NSI receives only statistical information as an output of the Solution and does not have access to the personal data itself. Nevertheless, it is the NSI who decides that the processing should take place, the processing is carried out for its purpose and its activity (producing official statistics) and it provides the MNO with detailed instructions on what information to collect. The NSI is thus to be considered a controller with respect to further processing pseudonymous mobile location data by means of the Solution for producing official statistics, which takes place in order to generate and deliver the final reports requested by the NSI. MNO may only process the mobile location data by means of the Solution for the purpose given by the NSI – for producing official statistics – according to the NSI's detailed instructions and is therefore to be regarded as processor.

At the same time, there are situations in which MNO may be considered to process mobile location data by means of the Solution also for its own purposes[399]:

1. **The MNO has discretion to accept or reject the offer to conclude an agreement for further processing pseudonymous mobile location data by means of the Solution for producing official statistics** – in those Member States where the national law has imposed a duty on the MNO to collect and process mobile location data for the purposes of enabling the NSI to produce official statistics from it, the MNO would be considered as controller with respect to the processing that is necessary to execute this obligation.[400] In other Member States, the MNO has discretion

[397] *Ibid*., sec 2.1.4, pp 14-15.
[398] *Ibid*., sec 2.1.5, pp 17.
[399] *Ibid*., sec 3.2.2., sec 55., p 20.
[400] *Ibid.*, sec 2.1.2, 1) sec 24, p 11.

to accept or reject the NSI's offer to conclude an agreement for further processing pseudonymous mobile location data by means of the Solution for producing official statistics – in these Member States, the controllership is established on the basis of an assessment of the factual circumstances surrounding the processing. In the latter case, the provisions of the agreement between the MNO and the NSI can be decisive for allocating (joint) controllership, provided that the agreement adequately reflects the real-world circumstances.[401]

2. **MNO-ND uses the Solution for generating pseudonymisation keys, which it then applies for pseudonymising the mobile location data** – since the pseudonymisation keys are generated randomly[402] and do not involve any personal data, such usage of the Solution by the MNO does not amount to processing of personal data for MNO's own purposes. Therefore, the issue of controllership does not arise because it presumes processing of personal data.
3. **MNO-VAD is provided with a copy of the final reports as output of the Solution** – the reason for making the final reports available to MNO-VAD is to help MNO-VAD ascertain the security of the processing operations by means of the Solution as well as to verify the anonymity of the mobile location data in the output of the Solution for MNO's internal regulatory and compliance purposes. Therefore, the MNO is not expected to gain benefits or influence the processing operations for its own purposes, especially considering that the final reports are ultimately intended to be made public by the NSI.

8.5. Lawfulness assessment

According to the DPIA Methodology, the check for lawfulness of processing has to be done prior to any DPIA.[403] Since only the Sample DPIA is conducted for the purposes of the Project, it is a suitable time to start considering potential legal basis for the actual statistical analysis use cases that will be developed in the future based on the Sample Use Case.

As further detailed in the process descriptions of the Sample Use Case (see Section 6.3.5 above), there are two general work processes involving personal data processing within the Solution:

a) **Pseudonymisation Process (P2)** – in order to better delineate the legal analysis below, it is presumed that the compatibility and lawfulness assessments for initial collection and pseudonymisation of the mobile location data have been conducted earlier and there is no need to further address them in this document. Therefore, it is presumed that the legal bases for the initial collection and pseudonymisation of the mobile location data are established elsewhere. The Sample DPIA only concerns further processing of such previously collected and pseudonymised mobile location data, even if the pseudonymisation process is carried out by means of the Solution.

[401] *Ibid.*, sec 2.1.2, 2) sec 28, pp 13-14.
[402] See: Solution Architecture. Description.
[403] *Op. cit.*, F. Bieker et al. A Process for Data Protection Impact Assessment. 2016, p 30.

b) **Application Work Process (P3)** – the legal analysis below is focused on the sub-process of the Application Work Process (P3), which implements the Sample Use Case (P5). More specifically, it considers potential legal bases for anonymisation conducted by means of the TEE component of the Solution (see Section 6.3.5.iii) for the purpose of producing official statistics2).

8.5.2. Potential existing legal bases under the EU statistics law

In this Section, a separate analysis is carried out to ensure a legal basis for anonymising mobile location data by means of the Solution for producing official statistics. As highlighted in the introduction to the Scoping Report, there is an uncertainty regarding applicable legal bases for NSIs to claim mobile location data from MNOs either in the EU law or in the relevant Member State's national law in order to reuse it for the purposes of official statistics. In order to resolve the uncertainty, there is a need to map potentially suitable legal bases for further processing pseudonymous mobile location data in established statistical production processes under currently applicable EU statistics law.

There are several established processes in the applicable EU statistics law, where mobile location data might be valuable as input for the production of official statistics:

1) **Demographic Statistics Regulation**
   - Art 3(1) requires Member States to provide the Commission (Eurostat) with data on their usually resident population at the reference time, including region of residence.
   - Art 4(1) requires Member States to provide the Commission (Eurostat) with data on the total population at national level at the reference time, for the purposes of qualified voting in the Council.
2) **Population and Housing Censuses Regulation**
   - Art 3 requires Member Sates to submit to the the Commission (Eurostat) data on the population covering determined demographic, social and economic characteristics of persons, families and households, as well as on housing at a national, regional and local level, as set out in the Annex.
   - The Annex lists the topics to be covered in Population and Housing Censuses, including obligatory topics depending on the geographical level (national level, NUTS 1, NUTS 2, NUTS 3, LAU 2 etc), for example the place of usual residence, total population, location of place of work, locality, location of living quarters etc.
3) **Persons and Households Regulation**
   - Art 3 requires the Member States to carry out data collection, inter alia, in the domain of income and living conditions.
   - According to Art 3(4)(e), the required data sets are further detailed in Annex I and involve duration of stay in the country, living environment, housing difficulties (including renting difficulties) and reasons etc.
4) **Migration Statistics Regulation**
   - Art 3 (1)(c)-(d) require Member States to supply the Commission (Eurostat) with statistics on the numbers on usually resident population.
5) **Tourism Statistics Regulation**

- Art 9 (1)-(3) requires Member States to transmit data, including confidential data, to Eurostat in the form of aggregate tables (e.g. breakdowns by duration and by destination of tourism trips for personal purposes) or micro-data files (with each observed trip being an individual record in the transmitted dataset, including month of departure, duration of the trip in number of nights, [only for outbound trips] duration of the trip: number of nights spent on the domestic territory, main country of destination etc).

Among the five established statistics production processes described above, there are only two potential directly applicable norms for producing statistics, which may serve as legal basis for NSIs to claim and process mobile location data under the EU law – Tourism Statistics Regulation Art 8 (b)[404] and Persons and Households Regulation Art 9(1). This is because neither of these two norms require basing statistics on other appropriate sources "in accordance with national laws and practices". In other cases, a specific legal basis in Member State law is required to produce statistics from mobile location data, because the condition "in accordance with national law and practice" is included in the relevant EU norm – Demographic Statistics Regulation Art 7 and Migration Statistics Regulation Art 9(1)(f).

Population and Housing Censuses Regulation Art 4(1) does mention the possibility to base statistics on sources other than those explicitly mentioned in the Regulation (conventional censuses, register-based censuses, sample surveys, rolling censuses). However, there is no specification as to which law those sources should be based on (no reference to "in accordance with national laws and practices"). Therefore, further processing pseudonymous mobile location data is questionable under that regulation.

To conclude, there might be a possibility to process mobile location data for the purposes of producing tourism statistics, persons and households statistics and population and housing censuses. However, this possibility needs to be checked against the GDPR, e.g. whether further processing pseudonymous mobile location data under the relevant regulations of statistics law complies with the GDPR Art 6(1)(e) requirements for processing personal data as part of a public interest task or exercising official authority. This analysis is out of scope of the Sample DPIA.

Another option is to consider legal bases for carrying out experimental statistics based on mobile location data, for example pilot and feasibility studies. Four out of the five established statistics production processes described above mention such studies:

- Demographic Statistics Regulation Art 8 and 11(1) – the study was to be carried out and the results to be delivered to the Commission by 31.12.2016, thus the legal basis is not relevant for the purposes of this analysis anymore.
- Persons and Households Regulation Art 14 – grants a right, but not an obligation, for Eurostat to launch and Member States to participate in feasibility and pilot studies for improving statistical methodologies upon necessity. Since the studies are of voluntary nature, it is questionable if this

[404] *Op. cit.*, G. Somers, 2017, p 51.

EU norm can be relied on as a legal basis for further processing pseudonymous mobile location data in terms of GDPR Art 6(1)(c) or (e).

- Migration Statistics Regulation Art 9a (1) – creates an obligation for Eurostat to establish pilot studies for improving the deployment of (new) data sources for migration statistics but leaves participation for Member States voluntary. Thus, just as in the previous case above, it is questionable if this EU norm can be relied on as a legal basis for further processing pseudonymous mobile location data in terms of GDPR Art 6(1)(c) or (e).
- Tourism Statistics Regulation Art 5 – creates an obligation for the Commission to draw up a programme for pilot studies for improving tourism statistics but leaves participation for Member States voluntary. Thus, again, it is questionable if this EU norm can be relied on as a legal basis for further processing pseudonymous mobile location data in terms of GDPR Art 6(1)(c) or (e).

In conclusion, there seems to be a possibility to carry out pilot and feasibility studies based on mobile location data, but only in very limited cases and only on voluntary basis, when involving the Member States and local MNOs. Again, this possibility needs to be checked against GDPR Art 6(1)(e) requirements for processing personal data as part of a public interest task or exercising official authority. This analysis is out of scope of the Sample DPIA.

To sum up, there do not seem to be a clear legal basis under the EU statistics law for further processing pseudonymous mobile location data for the purposes of producing official statistics as foreseen in the Sample Use Case. Therefore, other types of legal basis may be applicable, as analysed in the next section.

8.5.3. Potential other legal bases

In the context of DPD, the WP29 explained that the legal basis for anonymisation can be found in any of the grounds mentioned in DPD Art 7 (now the catalogue of legal grounds in GDPR Art 6), provided that the data quality requirements of DPD Art 6 (now principles relating to processing of personal data outlined in GDPR Art 5) are also met and with due regard to the specific circumstances and all the factors mentioned in the WP29 Opinion on Purpose Limitation.[405]

Based on the same logic, there are several options for choosing a legal basis for anonymisation from the catalogue of GDPR Art 6, each of which will be addressed separately below.

i. Consent of the data subject (GDPR Art 6(1)(a))

Just as in a typical statistical analysis process, the Subscriber may be asked for a consent to anonymise his/her mobile location data by means of the Solution for producing official statistics. The consent should be given to the entity who acts as the controller – most probably NSI but in some cases MNO (see Section 8.4 above). The MNO can also facilitate the NSI in obtaining the consent.

[405] *Op. cit.,* Article 29 Data Protection Working Party, Opinion 05/2014 on Anonymisation Techniques, 2014, p 7.

Consent can also be used as an additional safeguard to ensure that the further processing is compatible with the initial purposes for which the mobile location data were collected (see Section 8.3.2 above).

The key requirement of consent as legal basis is that it must be freely given, specific, informed and unambiguous in terms of indicating the Subscriber's wishes by which he or she signifies agreement to the processing of personal data relating to him or her. This can be done by a statement or by a clear affirmative action,[406] such as ticking a box when visiting an internet website, choosing technical settings for information society services or another statement or conduct which clearly indicates in this context the data subject's acceptance of the proposed processing of his or her personal data. Silence, pre-ticked boxes or inactivity do not constitute consent. "If the data subject's consent is to be given following a request by electronic means, the request must be clear, concise and not unnecessarily disruptive to the use of the service for which it is provided."[407] "A declaration of consent pre-formulated by the controller should be provided in an intelligible and easily accessible form, using clear and plain language and it should not contain unfair terms. For consent to be informed, the data subject should be aware at least of the identity of the controller and the purposes of the processing for which the personal data are intended. Consent should not be regarded as freely given if the data subject has no genuine or free choice or is unable to refuse or withdraw consent without detriment."[408]

Consent cannot be relied on as a valid legal ground for further processing pseudonymous mobile location data under GDPR Art 6(1)(a), if "there is a clear imbalance between the data subject and the controller, in particular where the controller is a public authority and it is therefore unlikely that consent was freely given in all the circumstances of that specific situation"[409]. In the context of further processing pseudonymous mobile location data by means of the Solution for producing official statistics, that may be the case – if the legal basis for anonymising mobile location data by means of the Solution for producing official statistics is constructed in a way that the Subscriber has no choice but to allow the processing, then consent may not be the suitable legal ground. This matter will require further analysis for each statistical analysis use case developed for real-world scenarios in the future.

ii. Controller's legal obligation (GDPR Art 6 (1)(c))

Both EU law and national law of the relevant Member State may impose a duty on the MNO to process mobile location data for the purposes of enabling the NSI to produce official statistics from it, assuming that the NSI has a corresponding legal basis to further process mobile location data (stemming from national or EU law). However, according to ePD Art 9(1), such mobile location data would have to be anonymised before the start of such processing.

[406] GDPR Art 4(11).
[407] GDPR Rec 32.
[408] GDPR Rec 42.
[409] GDPR Rec 43.

It reains to be seen if the relevant data protection authorities and courts accept the novel interpretations of the concept “made anonymous” in ePD Art 9(1) as proposed in the Sample DPIA in the context of privacy enhancing technology where no data is shared out of the data owner's organization. If yes, then this would mean that the “made anonymous” requirement is fulfilled in case of further processing pseudonymous mobile location data by means of the Solution for producing official statistics. If these interpretations are accepted, then any existing legal bases under national laws may be considered applicable, as long as processing of mobile location data is carried out by means of the Solution. If not, a new legal basis may need to be developed, so as to enable the Solution as an appropriate means for further processing pseudonymous mobile location data for producing official statistics.

In any case, the legitimacy of the relevant EU or national law will be a matter of further legal analysis for each statistical analysis use case developed for real-world scenarios in the future.

iii. Controller’s public interest task or exercise of official authority vested in the controller (GDPR Art 6(1)(e))

Both EU law and national law of the relevant Member State may impose a public interest task or official authority on the NSI to further process mobile location data for the purposes of producing official statistics under GDPR Art 6(1)(e), assuming that the MNO has a corresponding legal basis to process mobile location data for the purposes of enabling the NSI to fulfil its task or authority (stemming from national or EU law) (see Section 8.5.3.ii above).

It should be noted, however, that as long as the legal basis leaves MNO some room of discretion whether to process the mobile location data or not (whether to enter into agreement with the NSI or not), the NSI cannot be expected to fulfil its task or authority as an obligation but rather as a discretionary decision. Therefore, two situations should be differentiated:

1. **voluntary data provision** – can rely on consent, e.g. in preparatory stages of a typical statistical analysis process where responses to survey questions are voluntary;[410]
2. **obligatory data provision** – requires the processing operations to follow from an explicit legal obligation[411] or, in its absence, processing operations are necessary for a performance of a task carried out in the public interest.[412]

For the purposes of the Sample DPIA, this differentiation is relevant when choosing the legal basis for different stages of adopting the Solution. If the further processing of pseudonymous mobile location data by means of the Solution for producing official statistics is carried out in the proof-of-concept or pilot project stage, the latter of which falls into the sphere of experimental statistics, then the voluntary data provision model seems most feasible, as the relevant statistical analysis use case

[410] *Op cit.*, European Data Protection Supervisor. EDPS Opinion on safeguards and derogations under Article 89 GDPR in the context of a proposal for a Regulation on integrated farm statistics, 2017, p 11.
[411] *Ibid.*, p 12.
[412] *Ibid.*

along with the appropriate statistical methodologies are still being developed. However, since the development of statistics is also part of the public interest task of the relevant NSI, it may be possible to apply also the obligatory data provision model.

If the further processing of pseudonymous mobile location data by means of the Solution for producing official statistics is carried out in the production stage, which falls into the sphere of official statistics, then the obligatory data provision model should be preferred, in order to provide a clear legal framework and assure transparency towards the Subscribers with regard to potential impact on their fundamental rights and freedoms.

Whether to rely on the voluntary data provision model or obligatory data provision model, will be a matter of further legal analysis for each statistical analysis use case developed for real-world scenarios in the future.

iv. Legitimate interests of the controller (GDPR Art 6(1)(f))

Theoretically, it is also possible that the MNO may rely on its legitimate interest under GDPR Art 6(1)(f) when further processing pseudonymous mobile location data by means of the Solution for producing official statistics. WP29 has pointed out that DPD Art 6(1) e) and, *inter alia*, ePD Art 9(1), demonstrate the need to keep personal data "in a form which permits identification" for no longer than is necessary for the purposes of the collection or further processing[413]. According to WP29, DPD Art 6(1) e) made a strong point that personal data should be anonymised "by default", i.e. "if the controller wishes to retain such personal data once the purposes of the original or further processing have been achieved, anonymisation techniques should be used so as to irreversibly prevent identification."[414]

In its Opinion on Anonymisation Techniques, WP29 acknowledged mobile operators' legitimate interest as a legal ground for making anonymous of the contents of traffic data as soon as possible after their collection. It did so under DPD Art 7(f) (legitimate interest), because this is allowed under ePD Art 6[415] (MNO's obligation to make traffic data anonymous) as *lex specialis* in relation to DPD. Based on this example, it becomes a question if the MNO could anonymise mobile location data based on the legitimate interest ground provided under GDPR Art 6(1)(f)?

The answer is yes, if ePD Art(9)(1) allows it. According to ePD Art 9(1), "[w]here location data other than traffic data […] can be processed, such data may only be processed when they are made anonymous". MNO's right to process location data other than traffic data directly depends on the condition that the data is "made anonymous" before further processing. Therefore, it can be argued that ePD Art 9(1) provides for a legitimate interest ground for anonymisation in compliance with GDPR

[413] *Op. cit.,* Article 29 Data Protection Working Party, Opinion 05/2014 on Anonymisation Techniques, 2014, p 7.
[414] *Ibid.*
[415] *Ibid.*, p 8.

Art 6(1)(f). Alternatively, the legitimate interest ground could also derive from GDPR Rec 49, as proposed by G. Spindler and P. Schmechel:[416]

> *(49) The processing of personal data to the extent strictly necessary and proportionate for the purposes of ensuring network and information security, i.e. the ability of a network or an information system to resist, at a given level of confidence, accidental events or unlawful or malicious actions that compromise the availability, authenticity, integrity and confidentiality of stored or transmitted personal data, and the security of the related services offered by, or accessible via, those networks and systems, by public authorities, by computer emergency response teams (CERTs), computer security incident response teams (CSIRTs), by providers of electronic communications networks and services and by providers of security technologies and services, constitutes a legitimate interest of the data controller concerned. This could, for example, include preventing unauthorised access to electronic communications networks and malicious code distribution and stopping 'denial of service' attacks and damage to computer and electronic communication systems.*

The legitimate interest ground in GDPR Art 6(1)(f) reads as follows:

> *Processing shall be lawful only if and to the extent that at least one of the following applies:*
> *[…]*
> *(f) processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child.*

Therefore, if the legitimate interest ground under GDPR Art 6(1)(f) is considered to be applicable by the relevant data protection authorities and courts, the legitimate interests of the MNO should be weighed against the rights and interests of the Subscriber, considering, *inter alia*, that at least some of the Subscribers will be children. This weighing exercise should be carried out as a matter of further legal analysis for each statistical analysis use case developed for real-world scenarios in the future.

One idea to consider when evaluating the legitimate interest ground for the MNO is social responsibility duty or similar corporate responsibility engagement, whereby the MNO wishes to "give back" to the society. This way, the legitimate aim might even contribute also to the interests or fundamental rights and freedoms of the data subjects. Whether this idea is feasible in practice, needs to be further analysed for each statistical analysis use case developed for real-world scenarios in the future.

It is considered doubtful that further processing pseudonymous mobile location data by means of the Solution for producing official statistics can be based on legitimate

[416] *Op. cit*., G. Spindler, P. Schmechel. Personal Data and Encryption in the European General Data Protection Regulation, 2016, pp 173-174, sec 59-63.

interest of the MNO, considering that the NSI defines the purposes and the means of processing and, presumably, the actual legal basis for processing has to be connected to the purpose of the processing. It is difficult to envision how the MNO can rely on the legitimate interest ground solely on cyber security, social responsibility or data minimisation considerations, if the whole processing is carried out on the initiative of the NSI and fulfils a public purpose aim rather than commercial interests of the MNO. It shall remain to be clarified as a matter of further legal analysis for each statistical analysis use case developed for real-world scenarios in the future.

If, for some reason, the legitimate interest ground under GDPR Art 6(1)(f) does not apply, then it is still possible to rely on other grounds as discussed above (see Sections 8.5.3.i – 8.5.3.iii above).