PRACTICE

D31.3

Evaluation and integration and final report on legal aspects of data protection

Project number:	609611
Project acronym:	**PRACTICE**
Project title:	Privacy-Preserving Computation in the Cloud
Project Start Date:	1st November, 2013
Duration:	36 months
Programme:	FP7/2007-2013
Deliverable Type:	Report
Reference Number:	ICT-609611 / D31.3 / 1.0
Activity and WP:	Activity 3 / WP31.3
Due Date:	October 2016 - M36
Actual Submission Date:	7th November, 2016
Responsible Organisation:	UMIL
Editor:	Ernesto Damiani
Dissemination Level:	PU
Revision:	1.0
Abstract:	This deliverable contains the final report on the current legal framework regulating storage and processing the data on the cloud and complete the description of a methodology to analyze the business risks associated with outsourcing data, supported by a web-based tool.
Keywords:	Legal Framework, Secure computation, Data protection directive, Risk assessment methodology

This project has received funding from the European Unions Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 609611.

Editor

Ernesto Damiani (UMIL)

Contributors (ordered according to beneficiary numbers)

Stelvio Cimato (UMIL)
Ernesto Damiani (UMIL)
Gabriele Gianini (UMIL)
Francesco Viola (UMIL)
Jan Lundberg (UGOE)
Philipp Schmechel (UGOE)
Gerald Spindler (UGOE)

Executive Summary

The goal of PRACTICE's work package 31 is twofold: (i) reporting on the current legal framework regulating the protection of data stored and processed on the cloud and (ii) developing a risk assessment methodology for data sharing in cloud-based services.

In the first part of this deliverable, we provide a complete overview of the legal issues concerning cloud computing under the current European data protection law, including the new General Data Protection Regulation (GDPR), that will come in force in 2018. Our overview includes a detailed discussion of the most important regulations and directives, and the analysis of some case studies useful for evaluating the techniques developed by PRACTICE under a legal perspective. On the basis of our discussion, privacy and confidentiality breaches, especially the ones involving personal data, are identified as major regulatory issues at both European and national level.

In the second part of the deliverable, we complete the description of a process-oriented risk assessment methodology, aiming to analyze risks in multi-party business processes taking place on clouds. Our methodology supports the analysis and quantitative evaluation of risks related to privacy and confidentiality breaches during the execution of a collaborative process including the analysis of malicious coalitions of partners and the computation of the probability of information disclosure. It enables comparing risks of disclosure ”before” and ”after” the deployment of PRACTICE security controls. We also present a client-server version of our web-based tool supporting the modeling and the simulation of the business processes executed on the cloud

Chapter 1

Introduction

The objectives of work package 31 include the clarification of the legal framework which regulates the placing and the processing of sensitive data in locations where different privacy regulations hold, and the development of models and techniques to quantify the business risks associated with data sharing in collaborative services.
For this reason, D31.3 is divided into two parts. Part I is devoted to a complete overview of the current legal framework regulating data protection in the European Union, focusing on the legal challenges regarding cloud computing and encryption, and their relationship with the technologies developed by PRACTICE. In particular, the new issues caused by the adoption of the new General Data Protection Regulation, which will come into force on 25 May 2018, are examined discussing their relevance to the processing of personal data on the cloud, and the applicability of the Data Protection Directive to encrypted data, as well as the new obligations coming for controllers and processors (e.g. regarding informed consent of data subjects, transparency or data breach notifications, etc). Specifically, Chapter 2 provides a detailed analysis of the directive, focusing on the distinction of roles and responsibilities from the legal point of view between data controller and data processor, and outlining the impact of the new Data Protection regulation. In turn, Chapter 3 reports on some case studies where security controls implementing *Secure Multiparty Computation* (SMC) techniques have been deployed, and discusses their compliance with the outlined legal framework.

Part II contains the complete description of a risk assessment methodology that supports *comparative analysis* and *quantitative evaluation* of risks related to privacy and confidentiality breaches during the execution of any multi-party business process on the cloud. Specifically, it enables comparing risks of disclosure of a business process ”before” and ”after” the deployment of PRACTICE security controls and is supported by an open source web tool that can be used both to model and simulate the business processes executed on the cloud. In particular Chapter 4 focuses on the description of the process-oriented assessment methodology, while Chapter 5 discusses the coalition analysis of the collaborative system that can be used to quantify the likelihood of threats, showing a simple application example. A description of the architecture of the tool and a report on its usage is also contained in Chapter 6

Part I

Part I - Legal Status on Data Protection

Chapter 2

Cloud Computing under the European Data Protection Law

When Cloud Computing is used, legal problems might arise for every party involved. It can particularly be hard to achieve compliance with the data protection law. Chapter 2 of this deliverable provides an analysis of the European data protection law *de lege lata* and *de lege ferenda*.

2.1 Legal Framework

Prior to analysing specific problems for Cloud Computing arising from the data protection law, the basic functioning and key principles of the current and the upcoming state of law shall be outlined. The causes of and possible solutions for the legal difficulties can only be accurately explained and understood, provided that there is a general understanding of the underlying legal framework.

2.1.1 The Data Protection Directive 95/46/EC (DPD)

The Directive 95/46/EC on the protection of individuals with regard to the processing of personal data, and on the free movement of such data, was adopted in 1995 by the European Community to protect the privacy of individuals with regard to the processing of personal data.[1] EU Directives lay down certain end results that must be achieved in every Member State. National authorities have to adapt their laws to meet these goals and to implement the directives into their national law, but are free to decide how to do so. Nevertheless, directives have to be implemented in such a way that the best result is achieved ("effet utile"). Article 288 of the Treaty on the Functioning of the European Union defines how the Union's competences can be exercised. [2]

> "Article 288 (ex Article 249 TEC): [. . .] A directive shall be binding, as to the result to be achieved, upon each Member State to which it is addressed, but shall leave to the national authorities the choice of form and methods. [. . .]"

Each Directive specifies the date by which the national implementing laws must be adopted. A directive is addressed to the Member States, not to the citizens. Citizens may claim those rights directly only if directives state rights for citizens and if they are not implemented in due time by national authorities.

[1] 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, Official Journal of the European Communities, L 281, pp. 31-50.

[2] Treaty on the Functioning of the European Union, Official Journal C 326 of 26/10/2012, 0001 - 0390, available at: http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:12012E/TXT&from=EN.

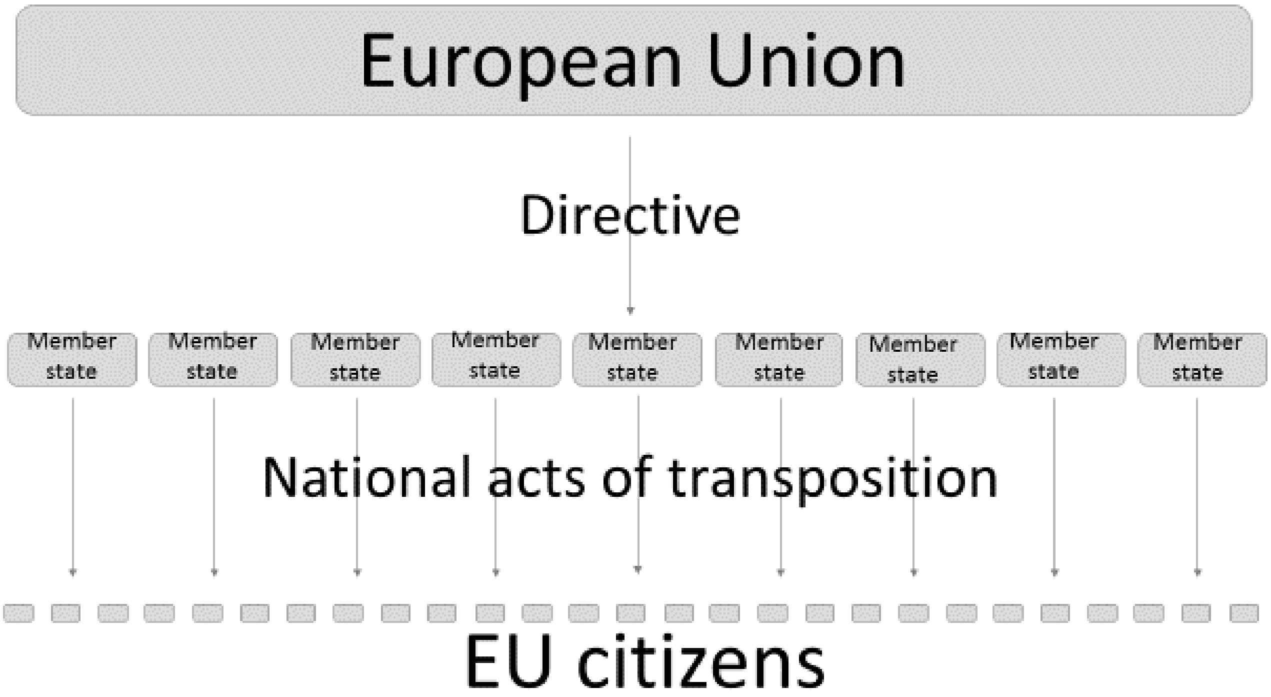


Figure 2.1: EU Directives

Directives are used to harmonize different national laws in order to create and foster the internal European market (e.g. product safety standards). [3]

Directives may differ concerning the grade of harmonization; for example, a *de minimis* harmonization allows Member States some leeway to pass laws which go beyond that level, but a full harmonization effectively prevents Member States from surpassing the directive.

Concerning the Directive 95/46/EC, the European Court of Justice (ECJ) passed a judgment in which it stated that the directive fully harmonizes the data protection law. This means the member states are not allowed to provide a lower level of protection than the directive demands, nor are they allowed to go beyond it. [4] Directive 95/46/EC imposes complete harmonization of national laws. [5] Directive 95/46/EC is intended to ensure free movement of personal data while guaranteeing a high level of protection for the rights and interests of data subjects, equivalent in all Member States. Consequently, Article 7 of Directive 95/46/EC sets out an exhaustive and restrictive list of cases in which the processing of personal data can be regarded as lawful. That interpretation is corroborated by the term “may be processed only if”, which demonstrates the exhaustive and restrictive nature of the list appearing in that article. Thus, the Member States can neither add new principles relating to the lawfulness of processing, nor impose additional requirements. [6]
Exempted from the scope of the directive (Article 3 Par. 2 of Directive) are areas related to the second and third so-called pillars of the European Union, i.e. the common foreign and security policy, police, and judicial cooperation in criminal matters.

[3] `https://europa.eu/european-union/law/legal-acts_en`.
[4] *ECJ*, decision of 24/11/2011 - C-J046/10.
[5] *ECJ*, decision of 24/11/2011- C468/10; *Kühling*, EuZW 2012, 281 (282).
[6] *ECJ*, decision of 04/10/2001 - C-450/00, Commission v Luxembourg.

The Directive generally prohibits the processing of personal data unless the person concerned has expressly consented to the processing of sensitive data or the processing is necessary to "keep the dissolution of the rights and obligations of the data controller in the field of employment law." In addition, the Directive allows Member States to provide for exceptions for reasons of substantial public interest.

In telecommunications, the data protection Directive is complemented by the regulation adopted in the 2002 Directive 2002/58/EC (Directive on privacy and electronic communications).

Recitals of Directives or Regulations do not have the same legal importance as provisions (Articles), but the Court of Justice of the European Union (ECJ) often refers to the Recitals when interpreting these provisions.[7]

2.1.1.1 Territorial Scope of the DPD

Since there might be a various number of parties (entities from all over the world) involved in cloud computing solutions, the important issue of international jurisdiction has to be addressed.

The DPD states that each Member State shall apply its data protection law when a "controller" carries out data processing by an establishment on the territory of a Member State. An exception to this principle is provided if the processor does not have an establishment in a Member State but uses equipment situated on the territory of a Member State for the purposes of processing. In this case, the European data protection law is applicable to the activities of the processor as well. Even an end-user's machine could be considered "equipment situated on the territory of a Member State" if it is used for storing a cookie or collecting data with java scripts. [8]

In contrast, if a webpage is accessible from the EU but hosted by a server in a third country, no equipment situated inside the EU is used. For the territorial scope of the Directive, it is not concerned with where a service is aimed at, but rather where the resources used for providing this service are located (this principle will change with the upcoming Data Protection Regulation, see 2.1.2.2). [9] A cloud server in Europe would qualify as "equipment" in the sense of the DPD. [10]

If a controller is established on the territory of several Member States, they have to ensure compliance with each of the applicable national laws, Article 4 DPD.

Even though Recital 19 of the DPD states that an establishment on the territory of a Member State "implies the effective and real exercise of activity through stable arrangements", there is no legal definition of "establishment" in the DPD. The ECJ clarified in a recent decision that the words "in the context of the activities of an establishment" cannot be interpreted restrictively. [11]

[7] *General Secretariat of the Council of the European Union.*, Manual of Precedents for Acts Established Within the Council of the European Union, 4th Ed., 2002, p. 83, available at: `http://bookshop.europa.eu/en/manual-of-precedents-for-acts-established-within-the-council-of-the-european-union-pbQC4101381/`; see also the references of the *ECJ* to Recitals of the DPD in Case C-131/12, Par. 48, 54, 58, 66, 67.

[8] As for example stated by the German court *KG Berlin* ZD 2014, 412 (414) regarding facebook's "friend finder".

[9] *Hon/Hörnle/Millard*, Data Protection Jurisdiction and Cloud Computing - When are Cloud Users and Providers Subject to EU Data Protection Law?, The Cloud of Unknowing, Part 3, p. 7 ff.; *Wieczorek*, DuD 2013, 644 (646); *Gabel*, in: Taeger/Gabel, BDSG, par. 1, recital 59.

[10] *Giedke*, Cloud Computing, p. 205 ff.

[11] *ECJ*, decision of 01/10/2015, Case C-230/14, Recital 25 = NJW 2015, 3636 (2368) - Weltimmo s. r. o. v Nemzeti Adatvédelmi és Információszabadság Hatóság, see also *Plitz*, K&R 2015, 559 (560); see also for a broad interpretation of "establishment" Art. 29-Working Party, Update of Opinion 8/2010, WP 179 update, 3.

According to the ECJ, the concept of "establishment" is versatile as it "departs from a formalistic approach whereby undertakings are established solely in the place where they are registered. Accordingly (...) both the degree of stability of the arrangements and the effective exercise of activities in that other Member State must be interpreted in the light of the specific nature of the economic activities and the provision of services concerned. This is particularly true for undertakings offering services exclusively over the Internet" (Recital 29). On the other hand, it is generally not necessary that the establishment should be independent from the controller in order to be considered a controller itself (for the definition of "data controller" see 2.3). [12]

Another case of the ECJ concerned a request for a preliminary ruling by the Supreme Court of Austria (OGH). The case, Verein für Konsumenteninformation v Amazon EU Sárl, refered to an action of the Austrian consumer association against Amazon, which has its European seat in Luxemburg. The case was concerned with Amazon's clauses in its standard terms and conditions according to which data from Austrian users might be exchanged with credit-risk assessment and financial services companies in Germany and Switzerland. The association had argued that Austrian data protection law should apply to this case. The question of the Austrian court to the ECJ was:

> "4.2. Is the processing of personal data by an undertaking that in the course of electronic commerce concludes contracts with consumers resident in other Member States, in accordance with Article 4(1)(a) of Directive 95/46/EC (...) and regardless of the law that otherwise applies, governed exclusively by the law of the Member State in which the establishment of the undertaking is situated in whose framework the processing takes place or must the undertaking also comply with the data protection rules of those Member States to which its commercial activities are directed?" [13]

The ECJ decided that "the processing of personal data carried out by an undertaking engaged in electronic commerce is governed by the law of the Member State to which that undertaking directs its activities, if it is shown that the undertaking carries out the data processing in question in the context of the activities of an establishment situated in that Member State". [14]
One of the cases decided by the ECJ highlights the difficulties to handle the notion of establishment in the DPD in practice. [15] The arguments brought forward by the Advocate General in the case of Google vs. Spain are worth being cited literally in order to emphasize the spectrum of interpretation concerning the notion of "establishment":

> "In my opinion the Court should approach the question of territorial applicability from the perspective of the business model of internet search engine service providers. This, as I have mentioned, normally relies on keyword advertising which is the source of income

[12] The German court Oberverwaltungsgericht (OVG = circuit court in administrative affairs) Schleswig-Holstein had to decide whether or not European data protection law was applicable for the data processing of Facebook, also in which European country Facebook respective establishment is acting. The court ruled that even though the US-American parent company Facebook Inc. is the only shareholder of the Irish subsidiary Facebook Ltd., the Irish company can be qualified as an establishment within the EU as Facebook Ireland obviously handled some of the data processing, OVG Schleswig Holstein, decision of 22/04/2013; however, another German court (Kammergericht (KG) Berlin (circuit court in civil law issue) in its ruling from 24/01/2014) contradicted that perspective that since the parent group Facebook Inc. is responsible for all decisions concerning data processing in the end, the Irish subsidiary Facebook Ltd. is not an establishment in the sense of the directive. This interpretation of 'establishment' does not comply with the directive's distinction between 'controller' and 'establishment'.

[13] Request for a preliminary ruling of the ECJ from the Oberster Gerichtshof (Austria) lodged on 27 April 2015, Case C-191/15 - Verein für Konsumenteninformation v Amazon EU Sárl.

[14] *ECJ*, decision of 28 July 2016 - Case C-191/15, Recital 81 - Verein für Konsumenteninformation v Amazon EU Sárl.

[15] *ECJ*, decision of 13/05/2014 - C-131/12 - Google Spain SL/Google Inc. v AEPD/Mario Costeja Gonzalez.

> and, as such, the economic raison d' être for the provision of a free information location tool in the form of a search engine. The entity in charge of keyword advertising (called 'referencing service provider' in the Court's case-law) is linked to the internet search engine. This entity needs presence on national advertising markets. For this reason Google has established subsidiaries in many Member States which clearly constitute establishments within the meaning of Article 4(1)(a) of the Directive. It also provides national web domains such as google.es or google.fi. The activity of the search engine takes this national diversification into account in various ways relating to the display of the search results because the normal financing model of keyword advertising follows the pay-per-click principle.
> 65. For these reasons I would adhere to the Article 29 Working Party's conclusion to the effect that the business model of an internet search engine service provider must be taken into account in the sense that its establishment plays a relevant role in the processing of personal data if it is linked to a service involved in selling targeted advertisement to inhabitants of that Member State.
> 66. Moreover, even if Article 4 of the Directive is based on a single concept of controller as regards its substantive provisions, I think that for the purposes of deciding on the preliminary issue of territorial applicability, an economic operator must be considered as a single unit, and thus, at this stage of analysis, not be dissected on the basis of its individual activities relating to processing of personal data or different groups of data subjects to which its activities relate.
> 67. In conclusion, processing of personal data takes place within the context of a controller's establishment if that establishment acts as the bridge for the referencing service to the advertising market of that Member State, even if the technical data processing operations are situated in other Member States or third countries." [16]

In the final judgment, the ECJ followed the Advocate General's opinion:

> "55 In the light of that objective of Directive 95/46 and of the wording of Article 4(1)(a), it must be held that the processing of personal data for the purposes of the service of a search engine such as Google Search, which is operated by an undertaking that has its seat in a third State but has an establishment in a Member State, is carried out 'in the context of the activities' of that establishment if the latter is intended to promote and sell, in that Member State, advertising space offered by the search engine which serves to make the service offered by that engine profitable." [17]

According to the Article 29 Data protection Working Party, in consequence of the ECJ's ruling in "Google Spain", even if a local establishment is not involved in any direct way in the processing of the EU data protection law may still be applicable, as long as there is an "inextricable link" between the activities of the local establishment and the data processing; thus, the "inextricable link" test has to be considered as a new element to the analysis of "in the context of the activities". [18]

[16] Opinion of Advocate General Jääskinen, delivered on 25/06/2013, Case C-131/12 - Google Spain SL/AEPD, recital 67.

[17] *ECJ*, decision of 13/05/2014, Case C-131/12 - Google Spain SL/Google Inc. v AEPD/Mario Costeja Gonzalez, rec. 55; c.f. the decision of the District Court of Heidelberg - *LG Heidelberg*, decision of 09/12/2014 - 2 O 162/113, recital 32 = MMR 2015, 348 (349), which decided that a link displayed by the Google search engine had an unambiguous reference to Germany and that thus German law was applicable.

[18] *Art. 29-Working Party*, Update of Opinion 8/2010, WP 179 update, 4 ff.

In the case of cloud computing, the DPD would be applicable if the cloud is an entity established within the jurisdiction of a Member State and if data processing is carried out within the context of the activities of this establishment. Every instance of processing a provider intends to carry out would then be covered, including the transfer of data to a non-EU country. In general, the directive is applicable if the cloud provider processes the data on a server within a Member State. If the provider is processing data using a machine physically in a certain Member State, this state's law is applicable as long as the provider is not established in another EU Member State. However, according to the above-mentioned ECJ decision "Google Spain", it is already sufficient that the Directive can be applied due to the fact that the cloud provider was established and is active in an EU Member State. It is not necessary that this establishment is directly involved in processing the data or has any particular responsibility concerning the processing; it is sufficient from an economic perspective that the establishment supports the activities of the cloud provider, such as in the Google Spain Case regarding the selling of advertisement etc. This however does not mean that the subsidiary company can be sued as the actual processor remains responsible. Hence, it is sufficient to apply the DPD provided that an establishment operates the funding for the cloud provider.

2.1.1.2 Material Scope and Fundamentals of the DPD

The focus of the Directive concerning "protection of individuals with regard to the processing of personal data and on the free movement of such data" (informal: "Data Protection Directive") is mentioned in Article 1:

"Object of the Directive
(1) In accordance with this Directive, Member States shall protect the fundamental rights and freedoms of natural persons and in particular their right to privacy with respect to the processing of personal data.
(2) Member States shall neither restrict nor prohibit the free flow of personal data between Member States for reasons connected with the protection afforded under paragraph 1."
[19]

Thus, the Directive again clarifies the two goals of fostering the internal market and guaranteeing basic rights for individuals concerning the protection of their personal data (privacy).
The Directive regulates the processing of personal data regardless of whether such processing is automated or not.

2.1.1.2.1 Material Scope of the Data Protection Directive

The Directive only protects the personal data of individuals, whilst corporate entities are excluded from the scope of the Directive. "Personal data" is any information relating to an identified or identifiable person, regardless of which aspects of the person the information may affect. Some examples are privacy issues, such as in private or job-related spheres, characteristics, skills of an employee, psychological characteristics or elements of someone's biography. [20]
Whilst the Directive applies in general for all kinds of processing data, there are still a number of distinctions made by the Directive. In the case of non-automatic processing the Directive only addresses data processing stored in a (physical) dossier. However, in this report we will deal solely with requirements for automatic processing of data due to the character of cloud computing. Moreover,

[19] Directive 95/46/EC of the European Parliament and of the Council, Official Journal L 281 of 23/11/1995, 0031 - 0050.

[20] *Dammann*, in: *Dammann/Simitis*, EG-Datenschutzrichtlinie, Art. 2, p.109.

Article 3 Par. 2 refers to some exceptions: [21]

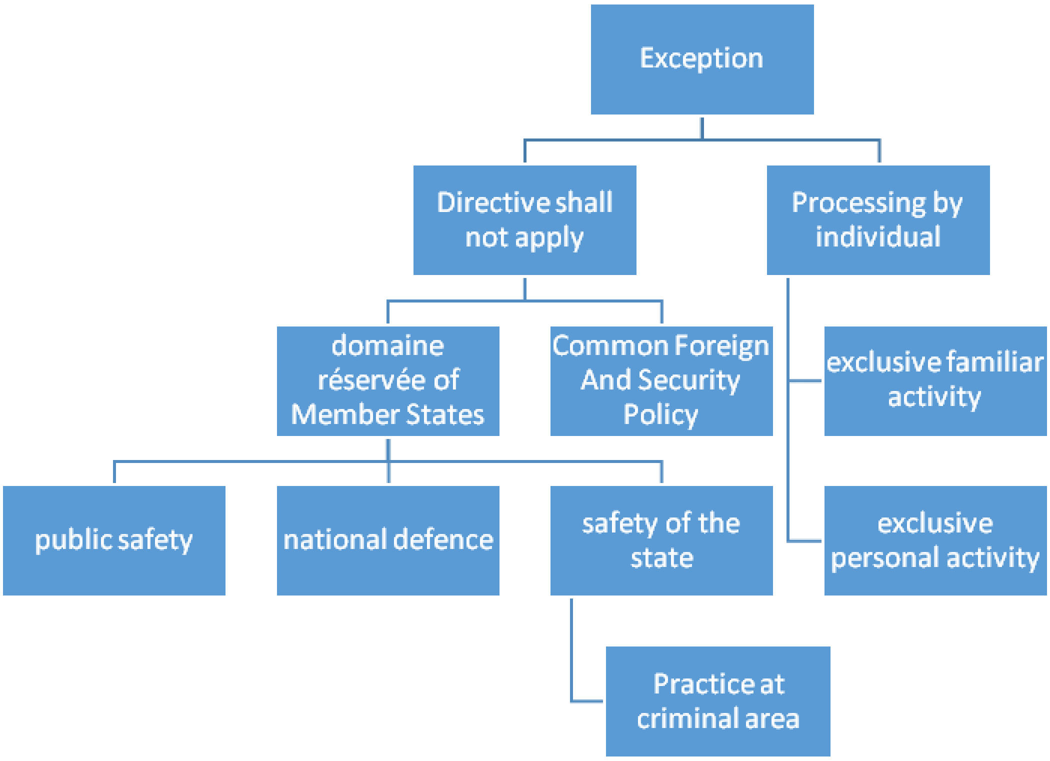

Exception
Directive shall not apply
Processing by individual
domaine réservée of Member States
Common Foreign And Security Policy
exclusive familiar activity
public safety
national defence
safety of the state
exclusive personal activity
Practice at criminal area

Figure 2.2: Applicability of the Data Protection Directive

One of the exceptions relevant for internet services (and users) refers to the exemption of exclusive personal and familiar activities. In other words, all activities on social networks etc. (user-generated content), which remain in the social and private sphere are not affected by the data protection directive. However, this exception does not alter the obligations of the operator of a social network.
Finally, the European Court clarified that processing for public safety and prosecution purposes is not within the scope of this Data Protection Directive.

2.1.1.2.2 Fundamentals of the Data Directive

The main principle is that personal data should not be processed at all unless the data processing operator complies with certain requirements. These refer to: transparency, legitimate purpose, and proportionality.

Transparency. The individual has the right to be informed should his personal data be processed, Article 10 and 11. Before starting the processing, the controller has to provide information about his identity (name and address), the purpose of processing, the recipient of the data and, if necessary, further information to guarantee fair processing in respect of the data subject. [22]
Personal Data can only be processed if the controller complies with the requirements stated in Article 7 and 12. Thus, explicit consent of the data subject is indispensable for the performance of contractual obligations or the entering into a contract.

[21] *Ehmann/Helfrich*, EG-Datenschutzrichtlinie, Art. 3, recital 16.

[22] Data subject is the official notion used by the Data Protection Directive, referring to the individual being affected by data processing.

Legitimate Purpose. Personal data shall only be "collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes", Article 6 (b).

Proportionality. Personal data may only be processed if the processing is "adequate, relevant and not excessive in relation to the purposes for which they are collected and/or further processed", Article 6. This processing has to be carried out "fairly and lawfully". Furthermore, the collected data must be "accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that data which are inaccurate or incomplete, having regard to the purposes for which they were collected or for which they are further processed, are erased or rectified".
Moreover, the directive demands the controller to "keep [the data] in a form which permits identification of data subjects for no longer than is necessary for the purposes of which the data were collected or for which they are further processed. Member States shall lay down appropriate safeguards for personal data stored for longer periods for historical, statistical or scientific use", Article 6 Par. 1e.
Finally, the directive tightens the requirements for specific sensitive personal data regarding "racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and [...] data concerning health or sex life". The processing of this kind of data may only be justified if the requirements stated in Article 8 Par. 2 are fulfilled such as a specific consent or protecting the vital interests of the data subject.

2.1.2 The General Data Protection Regulation

The General Data Protection Regulation (GDPR) [23] is a Regulation of the European Union, which will harmonize the rules regarding the processing of personal data mainly by private companies across the EU. The first proposal for a GDPR was published by the EU Commission on January 25th 2012. [24] On October 21st 2013, the European Parliament's LIBE Committee (Committee for Civil Liberties, Justice and Home Affairs) adopted a number of proposed changes to the General Data Protection Regulation, furthermore, on March 12th 2014, the European Parliament adopted a legislative resolution based on the proposal after the first reading in the parliament, adopting the LIBE Committee's changes to the original proposal. [25] On June 15th 2015, the Council of the European Union presented a new proposal for the General Data Protection Regulation with several changes and amendments[26], which led to the beginning of a series of intense trilogue-negotiations with the publication of the final - triloque - version of the GDPR in December 2015. The GDPR has now been passed and will finally

[23] 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), Official Journal of the European Union, L 119, pp. 1-88.

[24] Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) in the version adopted by the European Parliament after the LIBE-Committee's vote, available at: `http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+REPORT+A7-2013-0402+0+DOC+XML+V0//EN`; *Härting*, CR 2013, 715 (715 ff.).

[25] European Parliament, legislative resolution of 12/03/2014 on the proposal for a regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation), COM (2012) 0011-C7-0025/2012-2012/0011 (COD) (Ordinary legislative procedure: first reading), available at: `http://www.europarl.europa.eu/sides/getDoc.do?type=TA&reference=P7-TA-2014-0212&language=EN`.

[26] Council of the European Union, Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) of 15/06/2015 - ST 9565 2015 INIT, available at: `http://data.consilium.europa.eu/doc/document/ST-9565-2015-INIT/en/pdf`.. See for an overview of the legislative process of the GDPR *Albrecht*, Computer Law Review International 2016, pp. 33 ff.

come into force from 25 May 2018 (see Article 99 Par. 2 GDPR) and will, according to Article 94 Par. 1 GDPR, repeal the DPD. [27] The outdated[28] DPD from 1995 was intended to encourage the free movement of personal data within Europe by harmonizing national provisions on data protection. [29] However, the scope of implementation of the Directive led to different interpretations of the national data protection laws and to a minimum standard. [30] Hence, the Regulation is to ensure a uniform standard of data protection,[31] the protection of personal data and the free movement of such data within the European Union. However, the GDPR includes several opening clauses in which the Member States are given room for concretizations of articles of the Regulation or for entire new national laws. [32] Thus, in several parts of the European data protection framework there will still be no complete harmonisation within the EU.

2.1.2.1 Difference between a Directive and a Regulation

Regulations are passed either jointly by the EU Council and European Parliament, or by the Commission alone [33] and are the most direct form of EU law - as soon as they are passed, they have binding legal force throughout every Member State, with the same effects as national laws and eventually overruling them. National governments do not have to take action themselves to implement EU regulations. Art 288 of the Treaty on the Functioning of the European Union defines a regulation as:

> “Article 288 (ex Article 249 TEC): [. . .] A regulation shall have general application. It shall be binding in its entirety and directly applicable in all Member States. [. . .]”

The GDPR will therefore be directly binding and applicable in all Member States of the European Union without being in need of a national act of transposition.[34] This is an important difference between the current Directive and the Regulation, since the Directive had to be implemented into the national laws by the governments of Member States.

2.1.2.2 Territorial Scope of the GDPR

The GDPR uses a very broad territorial scope, especially the rules for controllers not established in the EU will be changed dramatically.[35] The territorial scope of the Regulation is specified in three cases, Article 3 Par. 1 - 3:

> “1. This Regulation applies to the processing of personal data in the context of the activities of an establishment of a controller or a processor in the Union, regardless of whether

[27] See for an overview of the legislative process of the GDPR *Albrecht*, Computer Law Review International 2016, pp. 33 ff.

[28] *Tene*, International Data Privacy Law 2011, 15 (15); *Hon/Millard*, Data Export in Cloud Computing - How can Personal Data be Transferred outside the EEA?, The Cloud of Unknowing, Part 4, p. 2; *Sartor*, International Data Privacy Law 2013, 3 (3).

[29] *Hon/Millard/Walden*, The Problem of ’Personal Data’ in Cloud Computing - What Information is Regulated?, The Cloud of Unknowing, Part. 1, p. 4; *Leonard*, International Data Privacy Law, 2014, 53 (53).

[30] *Klar*, ZD 2013, 109 (109 ff.); While one could have understood the Lindqvist- decision of the ECJ (of 06/11/2003 - C-101/91) in the way, that the Directive 95/46/EC requires only minimum standards of the Member States, it is obviously after the ASNEF-decision (24/11/2011- C-468/10), that the conditions of admissibility of the data handling are already largely fully harmonized.

[31] *Eckhardt/Kramer/Mester*, DuD 2013, 623 (630).

[32] See for an overview of the GDPR’s opening clauses `https://www.flickr.com/photos/winfried-veil/29706462112/in/datetaken-public/`.

[33] *Wieczorek*, DuD 2013, 644 (646).

[34] *Reding*, International Data Privacy Law 2012, p. 119 (121).

[35] See *Kindt*, CiTiP Working Paper 26/2016, pp. 13 ff.

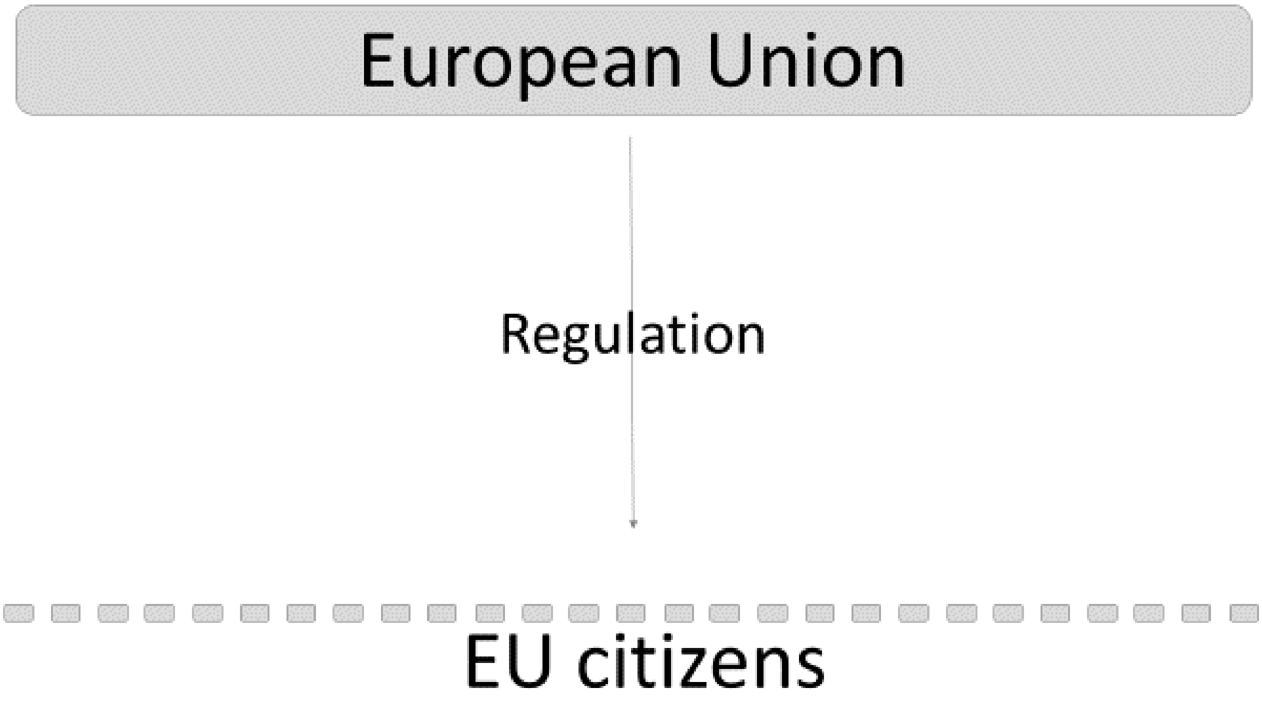


Figure 2.3: Applicability of the Data Protection Directive

> the processing takes place in the Union or not.
> 2. This Regulation applies to the processing of personal data of data subjects who are in the Union by a controller or processor not established in the Union, where the processing activities are related to:
> (a) the offering of goods or services, irrespective of whether a payment of the data subject is required, to such data subjects in the Union; or
> (b) the monitoring of their behaviour as far as their behaviour takes place within the Union. …"

Hence, many data processing operations by providers of services outside the European Union will thus fall into the scope of the European data protection law. The Recitals 22 and 23 highlight these intentions:

> "(22) Any processing of personal data in the context of the activities of an establishment of a controller or a processor in the Union should be carried out in accordance with this Regulation, regardless of whether the processing itself takes place within the Union or not. Establishment implies the effective and real exercise of activity through stable arrangements. The legal form of such arrangements, whether through a branch or a subsidiary with a legal personality, is not the determining factor in this respect.
>
> (23) In order to ensure that natural persons are not deprived of the protection to which they are entitled under this Regulation, the processing of personal data of data subjects who are in the Union by a controller or a processor not established in the Union should be subject to this Regulation where the processing activities are related to offering goods or services to such data subjects irrespective of whether connected to a payment."

Recital 23 GDPR continues by stipulating that it shall be apparent that the controller or processor envisages offering services to data subjects in the Union and that "the mere accessibility of the con-

troller's, processor's or an intermediary's website in the Union, of an email address or of other contact details, or the use of a language generally used in the third country where the controller is established, is insufficient to ascertain such intention". However, "factors such as the use of a language or a currency generally used in one or more Member States with the possibility of ordering goods and services in that other language, or the mentioning of customers or users who are in the Union, may make it apparent that the controller envisages offering goods or services to data subjects in the Union."
The concept of services is governed by Article 57 TFEU (freedom to provide services) or by Article 4 No. 1 of the Services Directive 2006/123/EC.[36] Services are all activities covered under Article 57 TFEU, which are normally provided for remuneration, insofar as they are not subject to the rules on free movement of goods, capital and on the free movement of the person. By making it clear in the definition of the Regulation that the service does not have to be paid for, commercial and non-commercial websites are covered. The definition of goods is governed by Article 28 Par. 2 TFEU. Regardless of the nature of the transactions, this is a set of objects which can be, in respect of commercial transactions, brought across a boundary. [37] These goods do not need to be physical, but must have a market value.
When the processing operation of the observation of the behavior of the person affected occurs, according to Recital 24 the Article 3 Par. 2 lit. (b) GDPR applies.

> "(24) The processing of personal data of data subjects who are in the Union by a controller or processor not established in the Union should also be subject to this Regulation when it is related to the monitoring of the behaviour of such data subjects in so far as their behaviour takes place within the Union. In order to determine whether a processing activity can be considered to monitor the behaviour of data subjects, it should be ascertained whether natural persons are tracked on the internet including potential subsequent use of personal data processing techniques which consist of profiling a natural person, particularly in order to take decisions concerning her or him or for analysing or predicting her or his personal preferences, behaviours and attitudes."

For example, when internet activities are tracked by means of data processing techniques by which a person is assigned to a profile, tracking-tools which operate on the use of cookies[38] for targeted advertising are particularly affected. [39] The Regulation applies to the processing of personal data by a controller not established in the Union, but in a place where the national law of a Member State applies by virtue of public international law according to Article 3 Par. 3 GDPR. Pursuant to Recital 25, this affects diplomatic spaces such as embassies or consulates. [40]
Hence, the former territorial principle of Article 4 of the DPD has been abandoned in favour of a more market- and user-orientated model. [41] This very broad territorial scope of the GDPR has the potential to strengthen the protection of European citizens' rights, since the provider of services or goods is bound to European data protection law irrespective where they are established. For cloud computing this might lead to two different outcomes, depending on how many parties are involved. If a cloud provider from a non-EU/EEA country offers their services directly to the data subjects

[36] *Wieczorek*, DuD 2013, 644 (647); Klar, ZD 2013, 109 (113); Treaty on the Functioning of the European Union, available at: `http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:12012E/TXT&from=EN`; Directive 2006/123/EC of the European Parliament and of the Council of 12/12/2006 on services in the internal market, available at: `http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32006L0123&from=EN`.

[37] *ECJ*, decision of 09/07/1992 - C-2/90, Recital 26.

[38] *Art. 29-Working Party*, Opinion 04/2012, WP 194, 1 ff.

[39] *Peifer*, K&R 2011, 2011, 692 (692 ff.); *Klar*, ZD 2013, 109 (113).

[40] *Art. 29-Working Party*, Opinion 08/2010, WP 179, 22 ff.; *Wieczorek*, DuD 2013, 644 (648).

[41] *Härting*, BB 2012, 459 (462); *Dammann*, ZD 2016, 307 (309).

inside the EU/EEA in a business-to-consumer-relationship, they will be governed by European data protection law. However, if the cloud provider offers their cloud services to a company in a business-to-business-relationship which uses the services to "process" its customers' data, the cloud provider is not considered to offer services or goods directly to the data subjects (the company's customers). [42]
Due to the GDPR's broad claim of applicability and the fact that citizens no longer have to consider the location of the processor's servers[43], the affected person might be more successful in regards to asserting his or her rights. [44]
However, this approach may go significantly beyond what could be considered realistically enforceable: A researcher established outside the Union could monitor - among others - EU-citizens' internet activities (even if their website is not even supposed to target EU-citizens), and therefore be governed by European data protection law - without even being aware of it. [45] European supervisory authorities are not able to act outside the Union. Article 51 Par. 1 GDPR only states that:

> "Each supervisory authority shall be competent for the performance of the tasks assigned to and the exercise of the powers conferred on it in accordance with this Regulation on the territory of its own Member State."

Thus afr, there is no solution to this problem. [46] Although Article 27 GDPR states that a controller or processor outside the Union that is affected by its data protection law shall designate in writing a representative in the Union, there are no possibilities for sanctions or measures against such controllers in the GDPR (apart from fines according to Article 84 Par. 4 lit. (a) GDPR if a controller does not designate a representative).[47]
Thus, the GDPR's broad territorial scope leads towards a new awareness of data controllers (also established outside the Union) regarding their processing of personal data. Therefore, technologies which minimise the use of personal data - especially encryption - and which avoid the application of the GDPR become even more important.

2.1.2.3 Material Scope of the GDPR

The General Data Protection Regulation does not abandon the basic principles of the DPD. It does not significantly change the concept of personal data. Just like the DPD, the GDPR follows a "black/white approach", hence the data are either personal or not, which means that if the data has a personal reference, all data protection rules apply and if not, it is outside the GDPR's scope. [48] According to Article 2 Par. 1 GDPR

> "This Regulation applies to the processing of personal data wholly or partly by automated means and to the processing other than by automated means of personal data which form part of a filing system or are intended to form part of a filing system."

Article 4 No. 1 S. 1 GDPR defines that "'personal data' means any information relating to an identified or identifiable natural person[49] (data subject')" which is the same wording as Article 2 (a) DPD. When

[42] *Hornung/Sädtler*, CR 2012, 638 (640).
[43] *Nebel/Richter*, ZD 2012, 407 (410).
[44] *Roßnagel/Richter/Nebel*, ZD 2013, 103 (104).
[45] *Spindler*, GRUR 2013, 996 (1003); *Spindler*, GRUR-Beilage 2014, 101 (107).
[46] *Hornung/Sädtler*, CR 2012, 638 (640).
[47] This has been criticized by the former German Federal Minister of Justice *Leutheusser-Schnarrenberger*, MMR 2012, 709 (710).
[48] *Forgó*, International Data Privacy Law 2015, p. 54 (59).
[49] Like in Article 1 Par. 1 of the DPD, the material scope of the GDPR only applies to the processing of personal data of *natural* persons according to Article 1 Par. 1 GDPR.

a reference to a natural person can be made the data protection principle of the GDPR always applies regardless of the data's content.[50]

The principle of prohibition with reservation of authorisation in the data protection law has not been weakened in the GDPR;[51] on the contrary, it has been enhanced pursuant to Article 6 GDPR.[52] The processing of personal data shall be, as regulated in the DPD, lawful only if the data subject has given consent in accordance with Article 7 GDPR (see 2.4.2.3.2) to the processing of their personal data or if after consideration the processing is necessary for legal purposes. Moreover, according to Article 2 Par. 2 lit. (c) GDPR, in general the applicability remains restricted to processing of personal data outside the private sphere. Thus, the Regulation does not apply to the processing of personal data by a natural person within the context of an activity that is purely personal or in the domain of the household, which implies that the actual cohabitation of the persons is affected and not their relationship in terms of family law matters.[53]

2.1.2.4 Fundamentals of the GDPR

As seen in 2.1.1.2.2 for the DPD, the GDPR also requires the controller to comply with the provisions regarding transparency, legitimate purpose and proportionality.

2.1.2.4.1 Transparency

Article 14 Par. 1 GDPR provides additional information regarding the data subjects in comparison to the transparency provisions of Articles 10 and 11 DPD. [54] It includes obligations to inform the data subject about, e.g. the identity of the data protection officer, the period for which the personal data will be stored, the existence of the right to request from the controller access to and rectification or erasure of the personal data, as well as the right to lodge a complaint to the supervisory authority (see in detail 2.4.2.3.2). Moreover, the GDPR extends the general information obligations in Article 33 to a specific communication to the data subject in case of a personal data breach (see 2.5.5). Furthermore, according to Article 30 GDPR each controller and processor shall maintain documentation of all processing operations under its responsibility.

2.1.2.4.2 Legitimate Purpose and other Fundamentals

Article 5 Par. 1 lit. (b) of the GDPR includes provisions regarding the purpose limitation, according to which personal data shall be "collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes." Article 5 Par. 1 lit. (c) stipulates that personal data shall be "adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ('data minimisation')". Further Processing of personal data for other purposes than the initial purposes for which the personal data have been collected may be lawful according to Article 6 Par. 4 GDPR by testing whether the processing for another purpose is compatible with the purpose for which the personal data are initially collected. [55]

In this regard, according to lit. (e) the existence of appropriate safeguards, which may include encryption or pseudonymisation has inter alia to be taken into account. Moreover, Article 5 Par. 1 lit.

[50] Cf. *Kranenborg*, in: *Peers/Hervey/Kenner/Ward* (eds.), The EU Charter of Fundamental Rights, 2014, Art 8, Recital 08.85; Art. 29-Working Party, Opinion 04/2007, WP 136, 6 ff.; *Karg*, DuD 2015, 520 (521).

[51] *Taeger* in: Taeger/Gabel, BDSG, 4a, Recital 4.

[52] *Härting*, CR 2013, 715 (717).

[53] *Dammann*, in: Simitis, BDSG, 1, Recital 243; *Gola/Lepperhoff* ZD 2016, 9 (10 ff.).

[54] *Härting*, Internetrecht, Recital 369.

[55] See regarding further processing and Big Data in the GDPR *Mayer-Schönberger/Padova*, The Columbia Science and Technology Law Review 2016, 315 (325 ff.).

(b) GDPR stipulates that "further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes ('purpose limitation')". According to Article 89 Par. 1 GDPR technical and organisational measures such as pseudonymisation shall ensure respect for the principle of data minimisation if processing is carried out for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes. At first sight, this could be interpreted as a relief for companies that run big data analyses, however, according to Article 14 Par. 1 lit. (c) and Par. 4 GDPR this would be very difficult in practice as the controller has to inform the data subject prior to further processing with information on other legitimate purposes as well as any additional relevant information.

However, the encryption technologies developed by PRACTICE could help controllers who further process personal data for other purposes than the initial purposes to comply with the GDPR.

Rules regarding "profiling" are set down in Article 22 GDPR, according to which "the data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her'.' According to Par. 2 it shall only be lawful if the processing is necessary for entering into, or for the performance of a contract, if it is expressly authorized by a Union or Member State law, or if it is based on the data subject's explicit consent. According to Par. 3 the controller shall implement suitable measures to safeguard the data subject's right to contest the decision.

Finally, Recital 48 GDPR provides an intra group exemption for companies by stating that "controllers that are part of a group of undertakings or institution affiliated to a central body may have a legitimate interest in transmitting personal data within the group of undertakings for internal administrative purposes, including the processing of clients' or employees' personal data." Thus, Recital 48 facilitates the exchange of personal data within a group of undertakings (defined in Article 4 No. 19 GDPR), however, the legitimate interests of the company still have to be weighed with the fundamental rights and freedoms of the data subject according to Article 6 Par. 1 lit. (f) GDPR and regarding the transfer of personal data to companies belonging to the group of undertakings but which are located in third countries the general principles for data transfer to third countries apply (Recital 48 S. 2 GDPR).[56]

2.2 Personal Data and Encryption

The European data protection law only applies if 'personal data' is processed. Because of that it is very important to understand what data qualifies as personal data. Depending on how 'personal data' is defined, the effect a valid encryption of this data takes, might be different.

2.2.1 Personal Data and Encryption in the DPD

In the following we assume that a data controller, i.e., cloud-computing client, holds information about data subjects and wants this information to be stored in a cloud computing environment. In the centre of any consideration concerning cloud-based information, processing is the definition of "personal data" provided by the Data Protection Directive (DPD). Information that is not, or ceases to be, "personal data", may be processed, in the cloud or otherwise, and is therefore not affected by data protection law requirements. Thus, if the information held by the data controller is considered to be personal data "in the cloud", in terms of data protection such cloud-computing operations (e.g. storing and processing in the cloud) would normally fall under the respective national data protection

[56] *Härting*, Datenschutz-Grundverordnung, 2016, Recital 490.

acts or within the scope of the DPD. In cloud computing, the "personal data" definitional issue is crucial with respect to *anonymized, pseudonymised* and encrypted data. Concerning encrypted data - be it encrypted while in transmission, storage or computations - one of the issues refers to the problem of whether it still qualifies as personal data.

2.2.1.1 Personal Data and Cloud Computing

As already outlined, the characteristics of personal data is crucial for the application of the Data Protection Directive. Therefore, we have to take a closer look on the criteria for assessing these characteristics:

2.2.1.2 Article 2 (a) Data Protection Directive

According to Article 2 (a) of the Data Protection Directive "personal data" shall mean any information relating to an identified or identifiable person ("data subject");[57] an identifiable person is one who can be identified, directly or indirectly, in particular by referencing an identification number, or to one or more factors specific to his physical, physiological, mental, economic, cultural, or social identity.
Whether or not information qualifies as "personal data" depends on the circumstances in each individual case. For instance, a common family name may not single someone out within a country but is likely to identify a pupil in a classroom. Moreover, if the data processing controller is able to combine information with other data in order to identify individuals, then the information that was originally considered "personal data" may change.

2.2.1.3 Recital 26 Data Protection Directive

Recital 26 of the Directive renders more precisely the notion of "personal data":

> *"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; 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."*

Thus, leaving aside apparently "non-personal" information. Indeed, Recital 26 of the DPD explicitly recognizes that information constituting "personal data" may be rendered "anonymous". Therefore the data can be used freely by data controllers/operators such as cloud computing operators, if it is being anonymized. Moreover, the transmission of data may fall outside of the scope of the DPD if the data no longer qualifies as personal data; otherwise, the data subject‘s consent is needed (see 2.4.2). This raises an important legal question: Is the cloud computing provider considered to be a data "processor" who processes personal data on behalf of the controller (see 2.3), i.e. the cloud computing client?
Unfortunately, Recital 26 of the DPD is prone to various interpretations and thus there is no clear answer: [58]

[57] *Kokott/Sobotta*, International Data Privacy Law 2013, 222 (223); CJEU, Joined Cases C-92/09 and C-93/09 Volker und Markus Schecke and Eifert [2010] ECR I-11063, par. 52, 53 and 87.

[58] *Hon/Millard/Walden*, The Problem of ‘Personal Data’ In Cloud Computing - What Information is Regulated? - *The Cloud of Unknowing*, Part 1, p. 13.

2.2.1.4 Relative or Absolute Identifiability of Persons

The criteria concerning the identifiability of persons required by Article 2 (a) DPD are still subject to debate; in particular, whether or not a so-called absolute or relative approach has to be the basis for assessing controller's abilities to identify a person. [59]
In short, the "absolute approach" determines all possibilities and chances in which the data controller would be able to identify the data subject individually. Thus, all ways and means for a data controller without any regard to expenses etc. are taken into account. Even theoretical chances of combining data so that the individual is identifiable are included. If identifiability is assessed absolutely, then it is sufficient for the application of personal data acts if anyone in the world is able to decrypt or decode the encrypted data. [60] Applied to cloud computing, as long as anyone in the world is able to decrypt the data set, the operations of the cloud computing provider are subject to data protection legislation, even if the cloud computing provider does not possess the key for decryption. Based on this approach data protection legislation is applicable, regardless of the applied encryption technique, as long as one entity holds the key for decoding.
In contrast, the "relative approach" considers the relevance of the necessary effort required by the data controller in order to identify the data subject. [61] Therefore, only realistic chances of combining data in order to identify an individual are taken into account. With regards to encryption issues, data protection legislation is only applicable if the data controller is able to decrypt a certain data set[62] - or, at least, has reasonable chances of obtaining the decrypting key.
In the case law of some courts, the trend is beginning to lean towards favouring a relative understanding[63] - in contrast to the opinion of some academics.[64]

2.2.1.4.1 The ECJ's approach

Despite its enormous practical impact, this aspect has not been completely clarified yet. However, in October 2014 the German Federal Court of Justice (BGH) requested the ECJ [65] for a preliminary ruling in accordance with Article 267 TFEU on the interpretation of the dispute regarding whether a dynamic IP address can be considered as personal data. [66] The German Federal Court of Justice

[59] *Bergt*, ZD 2015, 365 (365 ff.) provides an up-to-date summary of the different opinions.

[60] *Art. 29-Working Party*, Opinion 04/2007, 7; OLG Hamburg, MMR 2008, 687 (688); *Nink/Pohle*, MMR 2015, 563 (565) which criticize that consequently this approach would lead to the result that there would virtually be no more anonymous data; *Pahlen-Brandt*, DuD 2008, 34 (38).

[61] *Dammann* in: Simitis, BDSG, par. 3, Recital 32; *Gola/Klug/Körffer*, in: Gola/Schomerus, Bundesdatenschutzgesetz, 3, Recital 10; *Spindler*, Verhandlungen des 69. Deutschen Juristentages, Band I, Gutachten, 2012, F 115, 116; Schulz in: Beck'scher Kommentar zum Recht der Telemediendienste, 11 TMG, Recital 24; *Roßnagel/Scholz*, MMR 2000, 721 (723); *Meyerdierks*, MMR 2009, 8 (8 ff.); *Eckhardt*, K&R 2007, 601 (603); *Voigt*, MMR 2009, 377 (377); *Hon/Millard/Walden*, The Problem of 'Personal Data' In Cloud Computing - What Information is Regulated?, The Cloud of Unknowing, Part 1, p. 46.

[62] *Spindler*, Verhandlungen des 69. Deutschen Juristentages, Band I, Gutachten, 2012, F 115, 116.

[63] *England and Wales High Court* (Administrative Court), [2011] EWHC 1430 (Admin), Case No. CO/12544/2009, Recital 51 f.; *Upper Tribunal* (Administrative Appeals Chamber), [2011] UKUT 153 (AAC), Appeal Number: GI/150/2011, GI/151/2011, GI/152/2011, Recital 128; *House of Lords*, [2008] UKHL 47, recital 27; *The Paris Appeal Court,* decision of 15 May 2007 - Henri S. vs. SCPP; *Local Court of Munich,* decision of 30 September 2008 - 133 C 5677/08, Recital 26; *District Court of Wuppertal*, decision of 19 October 2010 - 25 Qs 10 Js 1977/08-177/10; *District Court of Berlin*, decision of 31 January 2013 - 57 S 87/08; different point of view: The Stockholm Lnsrtt, reference No. 593-2005, publication date 8 June 2005; *Local Court of Berlin-Mitte*, decision of 27 March 2007 - 5 C 314/06, Recital 20; *Administrative Court of Wiesbaden,* decision of 27 February 2009 - 6 K 1045/08.WI, Recitals 52 ff.

[64] *Brennscheidt*, Cloud Computing, p. 51; Kuner, European Data Protection Law, p. 92; *Pahlen-Brandt*, DuD 2008, 34 ff; Weichert, in: Däubler/Klebe/Wedde/Weichert, BDSG, 3, Recital 13, 15.

[65] *ECJ*, Case C-582/14 - Patrick Breyer v Bundesrepublik Deutschland.

[66] *German Federal Court of Justice (BGH)*, decision of 28/10/2014 - VI ZR 135/13 = MMR 2015, 131.

stated in its request to the ECJ that a relative approach could be in accordance with Recital 26 of the DPD, according to which "to determine whether a person is identifiable, account should be taken of all the means likely reasonably to be used to identify the said person" (Recital 25). [67] The court also stated in Recital 28 that the "wording of the provision of the directive appears to be ambiguous" and argued that even if means shall be taken into account that can be used by a third party to identify the person, a relative approach of the identifiability of the person affected would be possible if means are only taken into account that could realistically be used. However, the general opinion of the European Commission and of several Member States regarding this case tended to veer towards an absolute approach. [68]

Thus, the ECJ had to resolve the dispute between an absolute or relative approach regarding IP-addresses by interpreting Article 2 (a) DPD and especially Recital 26. [69]

On 12 May 2016 the Advocate General (AG), Campos Sánchez-Bordona published his opinion regarding this case, however, whilst the ECJ is not bound to follow his opinion, it often does so.[70] In his opinion, the AG contradicts an interpretation of "means likely reasonably to be used ... by any other person" in such a way that it would be sufficient that any third party might obtain additional data in order to identify a person[71] since this "overly strict interpretation would lead, in practice, to the classification as personal data of all kinds of information, no matter how insufficient it is in itself to facilitate the identification of a user".[72] Moreover, the AG emphasises that otherwise "it would never be possible to rule out, with absolute certainty, the possibility that there is no third party in possession of additional data which may be combined with that information".[73] This can be interpreted as a tendency of the AG towards a relative approach.

In its judgement of this case of 19 October 2016 the ECJ ruled that "(...) a dynamic IP address (...) is personal data within the meaning of that provision, in relation to that provider, where the [provider] has the legal means which enable it to identify the data subject with additional data which the internet service provider has about that person." [74]. The ECJ is reasoning its judgment as Recital 26 of the Directive refers to the means likely reasonably to be used by both the controller and by any other person'. Thus, its wording suggests, that for information to be treated as personal data' within the meaning of Article 2(a) of that directive, it is not required that all the information enabling the identification of the data subject must be in the hands of one person. [75] Since the provider is dependent on the collaboration of a third party (the internet service provider) to obtain the necessary information beneficial to identify the data subject, this can be a hint for an absolute approach of the ECJ. [76]

[67] C.f. *Brink/Eckhardt*, ZD 2015, 205 (209).

[68] *Bergt*, IP-Adressen: EU-Kommission gibt BGH Nachhilfe in Sachen Grundrechte, available at: `http://www.cr-online.de/blog/2015/09/13/ip-adressen-eu-kommission-gibt-bgh-nachhilfe-in-sachen-grundrechte/`.

[69] *German Federal Court of Justice (BGH)*, decision of 28/10/2014 - VI ZR 135/13 = MMR 2015, 131 (132 f.), recitals 27, 29 ff.

[70] Opinion of *Advocate General* Campos Sánchez-Bordona, delivered on 12 May 2016, Case C-582/14 - Patrick Breyer v Bundesrepublik Deutschland.

[71] Opinion of the Advocate General (see Note 70), Recital 64.

[72] Opinion of the *Advocate General* (see Note 70), Recital 65.

[73] Opinion of the *Advocate General* (see Note 70), Recital 65.

[74] *ECJ*, judgement of 19 October 2016, Case C-582/14 Patrick Breyer v Bundesrepublik Deutschland, Recital 49; cf. regarding the classification of dynamic IP addresses as personal data for access providers judged by the EJC, judgement of 24 November 2011, Case C70/10 Scarlet Extended SA v Sabam, Recital 51, which states that "[IP] addresses are protected personal data because they allow those users to be precisely identified."

[75] *ECJ*, Case C-582/14 (see Note 74), Recital 43.

[76] *Bergt*, Das Ende der Rechtssicherheit im Datenschutzrecht, available at:`http://www.cr-online.de/blog/2016/10/19/das-ende-der-rechtssicherheit-im-datenschutzrecht/`; different opinion Stadler, EuGH entscheidet zum Personenbezug von IP-Adressen, available at: `http://www.internet-law.de/2016/10/eugh-entscheidet-zum-personenbezug-von-ip-adressen.html#comments`.

However, the ECJ ruled that the possibility of combining a dynamic IP address with additional data held by an internet service provider does not constitute means likely reasonably to be used to identify the data subject 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. [77] By declaring the only legal means of the provider to be sufficient, the ECJ refers to a relative approach [78] and emphasizes that the crucial point is the legal possibility of the provider to enforce the third party releasing the data information.

To sum up, the decision of the ECJ can be interpreted as a rather relative approach in order to determine whether a person is identifiable. However, since the ECJ tends to combine both relative and absolute criterions, the identifiability of natural persons required by Article 2 (a) DPD will remain subject to future debates.

2.2.1.4.2 The Article 29 Data Protection Working Party's Approach

The position of the Article 29 Data Protection Working Party [79] describes its stance concerning Article 2 (a) DPD as follows:

> *" 'Anonymous data' in the sense of the Directive can be defined as any information relating to a natural person where the person cannot be identified, whether by the data controller or by any other person, taking account of all the means likely reasonably to be used either by the controller or by any other person to identify that individual. "Anonymized data" would therefore be anonymous data that previously referred to an identifiable person, but where that identification is no longer possible.*

Recital 26 also refers to this concept when it reads that:

> *"the principles of protection shall not apply to data rendered anonymous in such a way that the data subject is no longer identifiable". Again, the assessment of whether the data allows 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. This is particularly relevant in the case of statistical information, where despite the fact that the information may be presented as aggregated data, the original sample is not sufficiently large and other pieces of information may enable the identification of individuals."* [80]

This opinion of the Working Party is interpreted by some authors as being cryptic as the Working Party has used indirectly similar notions for other cases (mentioned in the opinion) which are implicitly favourable to the relative approach. [81] Others argue that the opinion includes a rather absolute stance. [82]

[77] ECJ, Case C-582/14 (see Note 74), Recital 46; see also in favour of an "unreasonableness" of using illegal means *Spindler/Nink* in: Spindler/Schuster (eds.), Recht der elektronischen Medien, 3rd Ed. 2015, 11 TMG Recital 8; *Brisch/Pieper*, CR 2015, 724 (728), who argue that the wording of "reason" is not compatible with the use of illegal means, but who are, however, against a strict classification of illegal means as unreasonable and thus recommend a consideration of each individual case.

[78] *Stadler* (see Note 76).

[79] `http://ec.europa.eu/justice/data-protection/article-29/index_en.htm`.

[80] *Art. 29-Working Party*, Opinion 04/2007, WP 136, 21; see also *Leonard*, International Data Privacy Law, 2014, 53.

[81] Cf. criticism of *Kühling/Klar*, NJW 2013, 3611 (3614); *Pahlen-Brandt*, DuD 2008, 34 f.

[82] Cf. *Eckhardt*, CR 2011, 339 (341, 343); *Stimerling/Hartung*, CR 2012, 60 (63).

Indeed, if one takes a closer look at the Working Party's statement, it should be noted that it considers not only the means potentially used by the controller to identify the data subject but also the means that might be used by third parties. [83] This instance can be regarded as an indication for an absolute approach. However, the Working Party seemingly recognizes situations in which a set of data should be regarded as personal data with respect to one entity but not with respect to another one,[84] which, in turn, implies a relative approach. The reason for this apparent contradiction is that the Working Party puts emphasis on the circumstances of the particular situation of the processing action rather than on the personal perspective (thus, whose capacities have to be considered: Only the ones of the controller or of any other person in the world?). Hence, on the one hand, the assessment of the data has to take into account means for identification that can be used by the controller or any other third party [85] in a concrete situation. Simply theoretical chances of identification are insufficient to constitute the personal characteristics of the data. [86]
The results of this opinion should, in many cases - especially with respect to encryption technologies - be similar to the relative approach, since both consider only realistic chances to identify the data subject.

2.2.1.4.3 Interpretations of "Identifiability" in different Member States

Singular indications of a relative approach can be found in the legislation of some EU Member States (in particular, Great Britain and Austria). The British Data Protection Act of 1998 expressly focusses in Part I, 1 on information that is - or is likely to come - in the possession of the data controller in order to assess the identifiability: [87]

> *" 'personal data' means data which relate to a living individual who can be identified - (a) from those data, or (b) from those data and other information which is in the possession of, or is likely to come into the possession of, the* **data controller***, and includes any expression of opinion about the individual and any indication of the intentions of the data controller or any other person in respect of the individual"*

This definition clearly differs from the provided formulations in Art. 2 (a) DPD and recital 26 of the DPD by taking (expressly) only the perspective of the controller. [88] One may note this instance while assessing British court decisions (such as the ones referred to above). Hence, the risk of inconformity with a Directive arises out of the provision in case the absolute approach prevails. Furthermore, within the EU Member States it does not appear to be the "usual" practice to implement the DPD requirements of the term "personal data" by expressly focusing on the controller's perspective only (which can be interpreted a sign for a relative understanding). [89] Therefore, a general stance of national legislators in the EU that are in favor of a relative approach to interpret the term "personal data" within the DPD cannot be determined from those single provisions.

A remarkable gradation was stated in the Austrian data protection law in par. 4 No. 1 DSG 2000: [90]

[83] *Art. 29-Working Party*, Opinion 04/2007, WP 136, 18 f.

[84] *Art. 29-Working Party*, Opinion 04/2007, WP 136, 15 f.

[85] Cf. also *Bygrave*, Data Privacy Law, p. 132.

[86] Art. 29-Working Party, Opinion 04/2007, WP 136, 15.

[87] Cf. *Kuner*, European Data Protection Law, p. 95 f.

[88] Cf. *Hon/Millard/Walden*, The Problem of 'Personal Data' In Cloud Computing - What Information is Regulated?, The Cloud of Unknowing, Part 1, p. 19, recital 97.

[89] Cf. List of provision formulations in *Kuner*, European Data Protection Law, p. 95 f.

[90] Austrian data protection act from 2000, BGBl. I Nr. 165/1999, last amendment 23/05/2013, BGBl. I Nr. 83/2013, available at: `https://www.ris.bka.gv.at/Dokumente/Erv/ERV_1999_1_165/ERV_1999_1_165.pdf`.

> “ ’Data’ (’Personal Data’): Information relating to data subjects (sub-par. 3) who are identified or identifiable; Data are “only **indirectly personal**” for a controller (sub-par. 4), a processor (sub-par. 5) or recipient of a transmission (sub-par. 12) when the Data relate to the subject in such a manner that **the controller**, processor or recipient of a transmission cannot establish the **identity** of the data subject by legal means”

The Austrian law ostensibly combines the relative and a rather absolute approach. With respect to the cited provision, data generally has to be rendered “personal” if the controller or any other person is capable of identifying the data subject (which indicates an absolute understanding). [91]
However, whenever the controllers themselves cannot identify the data subject by using lawful and reasonable means, all processing actions done by them are privileged in many provisions. [92] This special category is called “indirectly personal” data by Austrian law.

In other words, as long as identifiability can be denied on the basis of a relative approach, the Austrian data protection act is applicable, but with less strict requirements (with respect to the particular controller). For instance, the transmission of such “indirectly personal” data into third countries does not require a permission by the data protection authority (par. 12 section 2 No. 2 of the Austrian data protection act). Nevertheless it should be stressed that the DPD does not provide such a subcategory within the category of personal data; there is no differentiation between data that allow a direct identification of the data subject and those indirectly doing so. Both cases expressly constitute (one category of) personal data (see Article 2 (a) DPD). [93]
In order to avoid conflicts with the DPD (and constitutional law), there are trends to reduce the scope of the category of “indirectly personal” data in Austria by using a very restrictive interpretation of that term.[94]
Regarding a cloud computing scenario where encryption technology is used, the Austrian law could consider the processed information “indirectly personal” data relating to the controller, if the encryption has a sufficient level of security and the controller has at least no realistic chance to obtain the decryption-key (by lawful means). So the data would not fall outside the scope of data protection law completely, but only a reduced level of data protection provisions would be imposed by the controller. However, it is argued by some Austrian authors that even (securely) encrypted data does not render them “indirectly personal” - even though encryption might be a typical example for data from which a controller –who does not hold the decryption key - cannot identify the data subject. [95] As a consequence, even this kind of data would be (directly) personal data and hence the data protection act would apply comprehensively without any privilege. In sum, one should not generalize the Austrian law approach, since it is on the one hand based upon a unique interpretation of the DPD that assumes differences between a direct and an indirect identifiability, and is on the other hand subject to a controversy regarding the actual scope of the term “indirectly personal” data.

In this respect the second dispute concerns the technical demands to the level of encryption. Thus, the question is: which technical level of encryption has to be reached in order to assume that the reconstruction/decryption and *de-anonymization* of personal information/data is impossible - do we need absolute (*theoretical*) security or is *state-of-the-art* security sufficient? [96] To put it simply, the

[91] *Pollirer/Weiss/Knyrim*, Datenschutzgesetz 2000, par. 4, recital relating to Z 1, p. 20.
[92] *Pollirer/Weiss/Knyrim*, Datenschutzgesetz 2000, par. 4, recital relating to Z 1, p. 20. f.
[93] Cf. *Bergauer*, Jahrbuch Datenschutzrecht 2011, 55 (60).
[94] Cf. *Bergauer*, Jahrbuch Datenschutzrecht 2011, 55 (57 f.).
[95] Cf. *Bergauer*, Jahrbuch Datenschutzrecht 2011, 55 (62).
[96] This distinction is made, for instance, in nuclear law and other laws referring to “dangerous” technologies.

current question in the legal debate is: what level of *encryption* or *anonymization* must be achieved to avoid the applicability of the data protection law?

2.2.1.4.4 The Impact of the Absolute Approach upon Cloud Computing and Encryption

The absolute approach is a radical perspective that widens the scope of Data Protection. The DPD would not be applicable only if there is no theoretical chance for the cloud computing provider to re-combine the data in order to identify the data subject. In particular, encryption of data would not change the basic character of personal data itself, only render it more difficult for unauthorized people to access. [97]

Hence, from a radical stance, encryption makes it more difficult to identify and "read" the personal data. However, it does not exclude the theoretical outcome of obtaining a key and access to the data. Thus, from this perspective, encryption is considered more of a *technical security* measure to ensure that data is not accessible to unauthorized persons rather than changing the quality of data (in contrast to anonymizing it). Even if the encrypted data is being used, for instance in calculations, and if the data does not lose (in the decrypted version) the personal references, the DPD would be applicable, as the cloud computing provider would still theoretically be able to decrypt it. [98]

Alternatively, supposing that the data controller has key-coded/encrypted the original personal data (changed names to code numbers, with a "key" showing which number corresponds to which name), destroyed the original personal data, but still possesses the key, then individuals can be identified from the key-coded data when used in combination with the key. If encryption were applied to a data set, the whole data set would be transformed - not just names within the data set. However, where the data controller possesses the decryption key, encrypted personal data might be revealed in a similar way to key-coded data. If so, it would still be considered as "personal data".

2.2.1.4.5 The Impact of the Relative Approach on Cloud Computing and Encryption

As outlined above, the relative approach concentrates on the reasonable means for a data controller to identify the data subject and to get access to the personal data. If neither the cloud provider nor the cloud computing client (the data controller) keeps a master key to the respective data or data set of the customer, no personal data or information could be considered as processed or transferred abroad, since no-one could decrypt this data except the key holder. [99] Only the data subject who exclusively has the key could decode the data. Hence, consent to such processing or transfer is not required because no personal information would be implied during the process - neither at the very beginning, when the data is being transferred to the cloud computing client, nor afterwards, when transferred to the cloud computing provider. Furthermore, if only the cloud user holds the decryption key - and not the provider - the data cannot be rendered "personal", since the provider is not capable of decrypting the data and identifying the data subject. The user still processes personal data, since the assessment can vary depending on the particular person in question. [100]

In other words, the relative approach focuses on reasonable terms by which a provider may identify

[97] Orientierungshilfe - Cloud Computing, Version 2.0, p. 12, which states that '*regularly* data does not lose its character as personal data by encryption'; although, '*regularly*' means that there are special cases where encryption can have the effect, that no personal reference exists, c.f. *Eckhardt*, DuD 2015, 176 (179 f.).

[98] C.f. *Brennscheidt*, Cloud Computing, p. 52 f; *Nink/Pohle*, MMR 2015, 563 (566).

[99] *Hon/Millard/Walden*, The Problem of 'Personal Data' in Cloud Computing - What Information is Regulated?, The Cloud of Unknowing, Part. 1, p. 28; *Frauenhofer Institut*, Cloud-Computing für die öffentliche Verwaltung, p. 116; *Spies*, MMR-Aktuell 2011, 313727,

[100] Cf. *Hon/Millard/Walden*, The Problem of 'Personal Data' in Cloud Computing - What Information is Regulated?, The Cloud of Unknowing, Part. 1, p. 25.

the data subject, particularly if it would be economically (and legally [101]) feasible. [102] As possibilities/capacities of providers and their economic interest may vary widely, in general terms what qualifies as a reasonable effort to de-anonymize cannot be assessed. Thus, data is not "personal data" anymore - in the sense of the DPD - if the reference to individuals can, at least, under regular conditions, no longer be reconstructed, i.e., a decryption or de-anonymization is almost impossible. [103] If, under usual conditions, de-anonymization can be regarded as impossible, then the (anonymized) data cannot be qualified as 'personal'. [104]

Of course, the DPD can be applied if the encryption still implies personal information [105]; and in light of this one ought to bear in mind that any personal identifier (even an IP-address [106]) may be qualified as personal data, especially according to the ECJ's opinion (see above 2.2.1.4.1).

Offering no more than utility infrastructure services, IaaS (Infrastructure as a Service) and PaaS (Platform as a Service) providers (and certain Software as a Service providers) may not know whether the information being processed while using their services is really "personal data". Hence, some authors argue that it may even seem inappropriate to apply the DPD to such cloud infrastructure providers as the processing of personal data depends upon their customer's choices. [107]

With regard to future decryption tools, the relative approach concentrates on the actual available technologies - not on tools that will be available in the future. However, the actual technological capacities may change by that time, meaning that "identifiability" may change, as well. [108] Therefore, representatives of the relative approach [109] tend to apply the DPD to encrypted data because eventually technical tools may facilitate decryption, like the encryption of DVDs. [110] Thus, foreseeable technical developments should be taken into account when assessing the current quality of personal data. [111] In addition, the uncertainty of when decryption can be done reasonably should not be borne by the protected individual given the uncertainty of security levels provided by encryption. [112]

Nonetheless, in order to check if one falls into the scope of the DPD (if a new technology arises which had been unknown before) the data controller has to verify available technologies continuously; hence, a dynamic obligation is imposed upon the controller to regularly evaluate the used technologies. Concerning encryption technology as a means to change the character of "personal" data and render it "impersonal," the encryption operators have to continuously check the state-of-the-art encryption

[101] *Nink/Pohle*, MMR 2015, 563 (565); *Spindler/Nink*, in: Spindler/Schuster, Recht der elektronischen Medien, par. 11 TMG recital 8, 11; different opinion: *Bergt*, ZD 2015, 83 (85). For instance, non-disclosure provisions or secrecy legislation may impede any re-combination of data between different providers. The supporters of the absolute approach negate these barriers.

[102] *Gola/Klug/Körffer* , in: Gola/Schomerus, Bundesdatenschutzgesetz, par. 3, recital 44.

[103] *Gola/Klug/Körffer*, in: Gola/Schomerus, Bundesdatenschutzgesetz, par. 3, recital 44; Kroschwald, ZD 2014, 75, (78).

[104] *Kühling/Klar*, NJW 2013, 3611, (3613); *Dammann*, in: Simitis, BDSG, 3, Recital 32; *Gola/Klug/Körffe*r, in: Gola/Schomerus, Bundesdatenschutzgesetz, 3, Recital 10; *Polenz* in: Kilian/Heusser, Computerrechts-Handbuch, Part 13, Rechtsquellen und Grundbegriffe, Recital 59.

[105] *Spies* , MMR-Aktuell 2011, 313727

[106] C.f.*Bergt* , ZD 2015, 365 (370 f.); *Spindler/Nink*, in: Spindler/Schuster, Recht der elektronischen Medien, par. 11 TMG recital 11 ff; *ECJ*, Case C-582/14 - Patrick Breyer v Bundesrepublik Deutschland (see 2.2.1.3).

[107] *Hon/Millard/Walden*, The Problem of 'Personal Data' in Cloud Computing - What Information is Regulated?, The Cloud of Unknowing, Part 1, p. 26.

[108] *Art. 29-Working Party*, Opinion 04/2007, WP 136, 15; *Kroschwald*, ZD 2014, 75 (78).

[109] *Art. 29-Working Party*, Opinion 04/2007, WP 136, 7; *LG Frankenthal*, MMR 2008, 687 (689); *Hon/Millard/Walden*, The Problem of 'Personal Data' In Cloud Computing - What Information is Regulated?, The Cloud of Unknowing, Part 1, p. 14.

[110] *Kroschwald*, ZD 2014, 75 (79).

[111] *Spies*, MMR-Aktuell 2011, 313727.

[112] *Stadler*, Datenschutz: IP-Adressen als personenbezogene Daten.

technology. [113]

2.2.1.4.6 Conclusion

As the absolute approach extends the scope of the DPD to nearly all kinds of data processing, [114] from the perspective of the authors of this report (and from the perspective of the majority of authors), the stringer arguments tend to favor the relative approach. [115] Based on the absolute approach, data controllers (or data processors) cannot accurately assess if the DPD is applicable, since the DPD would be extended to an omnipresent law without any real boundaries. [116] Furthermore, it should be considered that the specific purpose of the Directive is the protection of the right to privacy of natural persons (see Article 2 No. 1 DPD). In scenarios where no realistic or "reasonable" chances to identify the data subject exist, with respect to the concrete situation of the processing actions, the purpose of the DPD is not affected at all. Therefore, it does not seem necessary to apply restricting data protection laws under those circumstances. [117]

Yet, we have to note that, even on the grounds of the relative approach, re-combinability of "harmless data" and creating profiles out of these data ('big data') do fall under the scope of the DPD. Even if at the beginning of data processing the data was not personal, we have to keep in mind, that every data processor has to check if the data they used is already "personal data" or not. [118] Furthermore, data which is related to things ("Internet of things") can turn out to be personal data if the data can be brought with reasonable effort [119] into a direct relationship with a person. However, the decision of the ECJ regarding the personal reference of dynamic IP addresses tends towards a relative approach, but includes several elements that could lead to a wide (absolute) interpretation of "identifiable".

2.2.1.4.7 Summary

As the previous paragraphs have illustrated, the technical requirements set forth by data protection laws concerning cloud computing and encryption - in particular, the standards - are still not fully settled. In a nutshell, based upon the required expenses, such as time and labour, encryption technologies must be in a way sophisticated that efforts to attribute information to persons (to decrypt) must be realistic.

According to the relative approach, the perspective of the data processor is relevant in order to assess the (un)reasonable efforts to decrypt the data, thus forgoing an objective point of view that would consider if anyone in the world would be able to decrypt it.

[113] *Jotzo*, Der Schutz personenbezogener Daten in der Cloud, p. 68; *Kroschwald*, ZD 2014, 75 (78 f.); *Art. 29-Working Party*, Opinion 04/2007, WP 136, 15; Cf. also *Roßnagel/Scholz*, MMR 2000, 721 (723).

[114] *Meyerdierks*, MMR 2009, 8 (10); *Peifer*, K&R 2011, 543 (544); *Spindler*, Verhandlungen des 69. Deutschen Juristentages, Band I, Gutachten, 2012, F 115.

[115] *Dammann* in: Simitis, BDSG, par. 3, recital 32; *Gola/Klug/Körffer*, in: Gola/Schomerus, Bundesdatenschutzgesetz, par. 3, recital 10; *Spindler*, Verhandlungen des 69. Deutschen Juristentages, Band I, Gutachten, 2012, F 115, 116; *Schulz* in: Beck'scher Kommentar zum Recht der Telemediendienste, par. 11 TMG, recital 24; *Roßnagel/Scholz*, MMR 2000, 721 (723); *Meyerdierks*, MMR 2009, 8 (8 ff.); *Eckhardt*, K&R 2007, 601 (603); *Voigt*, MMR 2009, 377; *Hon/Millard/Walden*, The Problem of 'Personal Data' In Cloud Computing - What Information is Regulated?, The Cloud of Unknowing, Part 1, p. 46.

[116] *Meyerdierks*, MMR 2009, 8 (10).

[117] Cf. *Eckhard*, CR 2011, 339 (342); *Härting*, ITRB 2009, 35 (37); *Maisch*, ITRB 2011, 13 (14).

[118] *Spindler*, Verhandlungen des 69. Deutschen Juristentages, Band I, Gutachten, 2012, F 116.

[119] *Gerlach*, CR 2013, 478 (479); *Spindler*, Verhandlungen des 69. Deutschen Juristentages, Band I, Gutachten, 2012, F 121; *Gola/Klug/Körffer*, in: Gola/Schomerus, Bundesdatenschutzgesetz, par. 3, recital 10.

2.2.2 Personal Data and Encryption under the GDPR

In this section we will examine the GDPR's provisions regarding encryption in the same cloud computing scenarios as outlined above for the DPD. [120] We will analyse the exact material scope of the GDPR and the Regulation's approach regarding the interpretation of the "identifiable natural person". We will determine whether encrypted data has to be regarded as anonymised or pseudonymised data and at which point encrypted data may lose its personal reference.

2.2.2.1 Encryption in the GDPR

Unlike the proposal of the Parliament[121], the final version of the GDPR does not provide a further definition of encrypted data, but mentions encryption in several provisions as a compliance requirement. According to Article 32 Par. 1 (a) GDPR, encryption is regarded as an appropriate technical and organisational measure to ensure the security of processing (see in detail 2.4.4.2).
Moreover, in case of a data breach, the controller is not required to communicate to the data subject if he or she has implemented encryption as a technical and organisational protection measure (Article 34 Par. 3 (a) GDPR). [122] Additionally, encryption is one of the "appropriate safeguards" of Article 6 Par. 4 (e) GDPR, which have to be taken into account when assessing the compatibility of a processing for a purpose other than that for which the personal data have been collected. [123] Finally, depending on the classification of encryption as pseudonymisation or not[124], the provisions of the GDPR regarding pseudonymous data[125] may be applicable for encrypted data, too.
The fact that there are regulations concerning encrypted data within the GDPR could be interpreted to mean that encryption does not prevent the applicability of the European data protection law: If encrypted data would not fall under the scope of the GDPR, regulations concerning encrypted data within the GDRP would make no sense entirely. This interpretation would support an absolute approach. However, it does not take into account that the qualification of data as personal or non-personal depends on the respective controller's evaluation. Hence, the regulations of the GDPR that concern encrypted data are interpreted merely as establishing rules for the controller. The GDPR's acknowledgment of encryption technologies and the benefit granted to the controller who encrypts data can offer an incentive to controllers to encrypt the affected person's data before processing it. However, it does not answer the question of whether or not encrypted data is considered personal data for a party that is unable to decrypt it. This remains dependent on the approach taken to define "identifiability" (see the following). Thus, it is apparent that this does not deal with the applicability of the GDPR, but rather with the protection of personal data. [126]

2.2.2.2 The "Identifiable Natural Person"

Article 4 No. 1 S. 2 GDPR introduces a new definition of the concept of an "identifiable natural person". According to this an

[120] See above 2.2.1.

[121] Article 4 No. 2b of the proposal of the European Parliament for a GDPR (LIBE proposal) defines encrypted data as "personal data, which through technological protection measures is rendered unintelligible to any person who is not authorised to access it", thus, according to LIBE, encrypted data shall just be a subcategory of personal data, which shall not lose its personal reference due to encryption.

[122] See 2.5.5.

[123] See 2.1.2.4.2.

[124] See 2.2.3.6.

[125] See 2.2.3.5.

[126] See 2.4.4.2 and Recital 83 GDPR for more details regarding these measures.

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

The definition distinguishes between identifiability on the basis of a reference to an identifier which can clearly identify a natural person, or due to special personal characteristics such as a person's sexual preferences or medical condition.[127] The GDPR utilises a broad approach regarding the interpretation of "identifiable natural person" however, some terms can also be interpreted in a relative way.
Recital 26 S. 3 GDPR states that

> "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."

On the one hand, the Recital refers to means reasonably likely to be used "by another person" which have to be taken into account, which veers towards an absolute approach, because this third person could be any person in the world. [128] This is also in tune with the scope of Article 8 of the Charter of Fundamental Rights of the EU (CFR), according to which "identifiable" has to be interpreted widely. [129] Considering the absolute approach, this would suggest that it is irrelevant whether the data has been encrypted or not: every piece of information that can be associated to a person must, therefore, be considered as personal data[130] - this would greatly extend the scope of the Regulation on the European level. [131] Hence, cloud services which store users' information would more likely fall under the scope of the regulation. On the other hand, it can be argued that that the extent of effort required to obtain the link between the affected person and the data should be greatly considered. [132]
Consequently, with a relative approach it can still be identified that the Recital may account for the means used by the respective controller and a third person - but only if those means are *reasonably likely* to be used. [133] If the data is not reasonably likely to be decrypted, the data could be considered non-personal (i.e. anonymous data) because the affected person would not be identifiable.
Moreover, stating in Article 4 No 1 S. 2 GDPR that every "identifier" shall contain personal references is another hint for a rather absolute approach of the Regulation regarding the identifiability of a natural person. [134] Additionally, Recital 26 GDPR states that using means for "singling out" the natural person directly or indirectly may make this person identifiable. Thus, a data subject may now be singled out for data processing even if it is unlikely that his or her name can be tied to the data, because even this could result in harming his or her privacy. [135]
On the other hand, the term "means reasonably likely to be used" suggests limitations through relative elements, in particular the notion of "reasonably". [136] Additionally, if a zero risk threshold

[127] Cf. *Härting*, Datenschutz-Grundverordnung, 2016, Recital 275 ff.

[128] Cf. *Zuiderveen Borgesius*, Computer Law & Security Review 2016, p. 256 (267) who interprets Recital 26 as "an absolute approach to identiability"; Polonetsky/Tene/Finch, Santa Clara Law Review, (Forthcoming) 2016, p. 593 (614).

[129] Cf. *Kranenborg*, in: Peers/Hervey/Kenner/Ward (eds.), The EU Charter of Fundamental Rights, 2014, Art 8, Recital 08.85.

[130] *Härting*, CR 2013, 715 (718).

[131] *Hullen*, PinG 2015, 210 (211).

[132] See also regarding the DPD the opinion of the Advocate General (see Note 85), Recital 68.

[133] *Lang*, K&R 2012, 145 (146).

[134] *Brink/Eckhardt*, ZD 2015, 205 (208); Buchner, DuD 2016, 155 ff.; *Härting Datenschutz-Grundverordnung*, 2016, Recital 279; different opinion: *Schantz*, NJW 2016, 1841 (1843).

[135] *Hon/Kosta/Millard/Stefanatou*, Tilburg Law School Legal Studies Research Paper Series No. 07/2014, 9; *Zuiderveen Borgesius*, Computer Law & Security Review 2016, 256 (267); *Marnau*, DuD 2016, 428 (430).

[136] Cf. *Esayas*, European Journal of Law and Technology, Vol 6, No 2 (2015), 6; *Härting Datenschutz-Grundverordnung*, 2016, Recital 282.

would be applied for any potential data user, no existing technique could achieve the required level of anonymisation. [137] Moreover, according to the Article 29 Data Protection Working Party (interpreting the DPD), "a mere hypothetical possibility to single out the individual is not enough to consider the person as 'identiable"'. [138] Recital 26 GDPR continues by stating objective factors which shall be relevant for the interpretation of the means used to identify a natural person:

> "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."

These factors illustrate a further attempt to limit the broad absolute elements of the GDPR's material scope. [139] Significant objective factors will be *inter alia* the state of science and technology, including future technological developments as well as the time and costs needed to identify somebody. [140] Since Article 2 (a) DPD and Article 4 No. 1 GDPR are very similar, the ECJ's decision (see 2.2.1.4.1) has a major influence on the general interpretation of defining "identifiability" also in the GDPR. [141] Data which does not have any personal references, for instance sheer machine data does not fall under the material scope of the GDPR. However, this data can still turn into personal data when related to a natural person, for instance in a big data or *Internet of things* scenario. [142]
Recital 30 of the GDPR now explicitly states that:

> "(n)atural persons may be associated with online identifiers provided by their devices, applications, tools and protocols, such as internet protocol addresses, cookie identifiers or other identifiers such as radio frequency identification tags. This may leave traces which, in particular when combined with unique identifiers and other information received by the servers, may be used to create profiles of the natural persons and identify them."

Thus, a lot of the data which originally had no personal reference will become personal data due to the association of online identifiers with natural persons. Additionally, natural persons can often be identified or be identifiable by "singling out"[143] their data. Thus, because of the broad material scope of the GDPR and of big data technologies, there are fewer and fewer possibilities to process data without a personal reference, in particular in the Internet of Things era.

2.2.2.3 Conclusion

The GDPR's material scope contains several parts which can be interpreted as relative approaches regarding the identifiability of natural persons, most prominently with the duty to include means only, if they are "reasonably likely to be used". Moreover, according to the ECJ, illegal means shall not be considered. Nevertheless, several other terms indicate a rather absolute approach of the GDPR, be it the wide scope of the online identifiers, the incorporation of "singling out" or that information obtained by a third person shall be sufficient to make the data personal for a controller. However, the GDPR's material scope will have to be interpreted widely, by using including a mix of both relative and absolute elements.

[137] *Esayas*, European Journal of Law and Technology, Vol 6, No 2 (2015), 6.

[138] *Art. 29-Working Party*, Opinion 04/2007, WP 136, 15; *Art. 29-Working Party*, Opinion 05/2014, WP 216, 8 ff.

[139] *Spindler*, DB 2016, 937 ff.

[140] *Härting Datenschutz-Grundverordnung*, 2016, Recital 284; *Zuiderveen Borgesius*, Computer Law & Security Review 2016, 256 (262).

[141] *Härting*, ITRB 2016, 36 ff.; *Keppeler*, CR 2016, 360 (364).

[142] Cf. *Karg*, DuD 2015, 520 (522); see more detailed in *Spindler/Schmechel*, JIPITEC 2016, 163 (168).

[143] Regarding singling out people without knowing their names (for behavioural targeting) see *Zuiderveen Borgesius*, Computer Law & Security Review 2016, 256 ff.

2.2.3 "Anonymous Information" and the GDPR

Although technologies to anonymise personal data are considered to be of high value to protect the fundamental privacy rights of the data subjects, the GDPR does not provide a specific article to regulate "anonymous information" in the Regulation, it is only mentioned in one Recital. According to Recital 26 S. 4 and 5 GDPR the:

> "principles of data protection should (...) 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. This Regulation does not therefore concern the processing of such anonymous information, including for statistical or research purposes."

Thus, as well as the DPD, the GDPR is not applicable to anonymous data. To examine whether data can be considered as anonymous; once again the problem of the identifiability of data subjects arises. In this regard, the possibility to anonymise personal data in the GDPR can be seen as another hint in favour of a relative approach, because given the possibilities to re-identify and combine data (big data), anonymous information could not be established when following a pure absolute approach. [144] Unfortunately, this "definition" in Recital 26 does not resolve the aforementioned dispute between the different approaches (relative vs. absolute) to define anonymisation. Moreover, there is legal uncertainty regarding the lawfulness of the anonymisation process, more precisely whether anonymising personal data means "further processing" of personal data. [145] The Working Party states in its WP 216 (regarding the DPD), that "anonymisation constitutes a further processing of personal data; as such, it must satisfy the requirement of compatibility by having regard to the legal grounds and circumstances of the further processing". [146] Nevertheless, this further processing of personal data is considered to be compatible with the original purposes of the processing but only if the anonymisation process leads to "reliable (...) anonymised information". [147] Furthermore, the data controller's legitimate interest always has to be balanced against the data subject's rights and fundamental freedoms. [148] Consequently, according to the Working Party, anonymisation can be compatible with the original purposes of the processing, but it would be a violation of data protection law if personal data was anonymised for purposes that are not compatible with the original purpose and if there were no other legitimate grounds for processing the data, such as the data subject's consent. [149] The Working Party clarifies this by providing as an example the anonymisation of the contents of traffic data immediately after its collection by mobile operators which performed deep packet inspection technologies. It was lawful in accordance with Article 7 (f) DPD, because of a legal permission stipulated in Article 6 Par. 1 of the e-Privacy Directive for certain traffic data which has to be erased or made anonymous as soon as possible when it is processed and stored by the provider of a public communications network or publicly available electronic communications service. [150]

Applying the Working Party"s interpretation to the GDPR, the Regulation's requirements regarding further processing need to be fulfilled when anonymising personal data. Thus, it has to be analysed whether anonymisation is a compatible use according to the GDPR, then no legal basis separate from

[144] *Härting, Datenschutz-Grundverordnung*, 2016, Recital 291.

[145] See *El Emam/Álvarez*, International Data Privacy Law 2015, p. 73 (79); *Hon/Kosta/Millard/Stefanatou*, Tilburg Law School Legal Studies Research Paper Series No. 07/2014, 12; *Esayas*, European Journal of Law and Technology, Vol 6, No 2 (2015), 4 ff.

[146] *Art. 29-Working Party*, Opinion 05/2014, WP 216, 3, 7.

[147] *Art. 29-Working Party*, Opinion 05/2014, WP 216, 7.

[148] Cf. *Art. 29-Working Party*, Opinion 05/2014, WP 216, 8.

[149] Cf. *Walden*, International Journal of Law and Information Technology 2002, 224 (233); *Esayas*, European Journal of Law and Technology, Vol 6, No 2 (2015), 4.

[150] *Art. 29-Working Party*, Opinion 05/2014, WP 216, 8.

that which allowed the collection of the personal data would be required (Cf. Recital 50 GDPR). Recital 49 GDPR states that:

> "(t)he processing of personal data to the extent strictly necessary and proportionate for the purposes of ensuring (...) information security (...) 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 (...)."

According to this, the anonymisation of personal data could be interpreted as necessary for ensuring information security and be, in accordance with Article 6 Par. 1 (f) GDPR, of legitimate interest to a controller. [151] Apart from this, according to Article 5 Par. 1 (b) GDPR a "further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89 (1), not be considered to be incompatible with the initial purposes". Furthermore, according to the compatibility test of Article 6 Par. 4 GDPR, account should be taken *inter alia* of the possible consequences of the intended further processing for data subjects. Since anonymisation, pseudonymisation and encryption are privacy preserving technologies[152], in most cases applying these tools on the data subject's personal data will be in their interest. However, the Working Party's opinion implies a non-existent weakness of the data protection law. Because as long as the data is anonymous, there is no threat to the privacy of the data subjects and as soon as a re-identification of the data is possible the GDPR with all its protective instruments is applicable again. Moreover, the need to justify the process of anonymisation itself could discourage the use of privacy-enhancing techniques. [153] However, with the use of Recital 49 GDPR, this dispute could possibly come to an end as soon as the GDPR comes into effect.

2.2.4 "Pseudonymisation" and "Pseudonymous Data" in the GDPR

Unlike in the DPD, the GDPR includes a definition of "pseudonymisation". According to Article 4 No. 5 GDPR, pseudonymisation

> "means the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person".

Moreover, "the application of pseudonymisation to personal data can reduce the risks to the data subjects concerned and help controllers and processors to meet their data-protection obligations" (Recital 28 S. 1 GDPR). Furthermore, Recital 28 S. 2 GDPR emphasises that the explicit introduction of "pseudonymisation" does not intend to preclude any other measures of data protection. Additionally, the wording of Recital 26 S. 2 GDPR states that "personal data which have undergone pseudonymisation, which could be attributed to a natural person by the use of additional information should be

[151] Cf. *Esayas*, European Journal of Law and Technology, Vol 6, No 2 (2015), 5; *Hon/Kosta/Millard/Stefanatou*, Tilburg Law School Legal Studies Research Paper Series No. 07/2014, 12, who criticise that this legitimate interest should also refer to processors.

[152] Cf. Recital 29 S. 1 GDPR which gives incentives for controllers to apply pseudonymisation when processing personal data; Article 5 Par. 1 (c) which regulates the principle of data minimisation, which is fulfilled by these technologies that reduce the amount of personal data.

[153] *Hon/Kosta/Millard/Stefanatou*, Tilburg Law School Legal Studies Research Paper Series No. 07/2014, 12; *Esayas*, European Journal of Law and Technology, Vol 6, No 2 (2015), 5.

considered to be information on an identifiable natural person". Thus, the connection between a natural person and the information on the basis of a corresponding rule remains - pseudonymised data is still qualified as personal data. [154]

Hence, pseudonymisation is merely a method which can reduce the likelihood of identifiability of individuals, but does not exclude this data from the material scope of the GDPR. It is handled by the Regulation primarily as a data security measure,[155] and its use is encouraged in several articles of the GDPR; Article 32 Par. 1 (a) names it an appropriate technical and organisational measures to ensure a level of security appropriate to the risk. Additionally, pseudonymisation is a technical and organisational measure that shall be implemented by the controller as a way to comply with the principle of data minimisation for the newly introduced provisions for "data protection by design and by default"(see 2.5.2). [156] Nevertheless, to clearly define the unclear provision and the use of pseudonymisation, associations and other bodies representing categories of controllers or processors may prepare "codes of conduct" according to Article 40 Par. 2 (d). [157]

2.2.5 Is Encrypted Data in the GDPR Anonymised or Pseudonymised?

Since the GDPR does not define "encrypted data", we have to examine if encryption is a technique which anonymises or just pseudonymises personal data. In this regard, again the dispute regarding the material scope of the Regulation (see 2.2.3.2) plays an important role.

When encrypting personal data, in accordance with Article 4 No. 5 GDPR, the encryption key is the "additional information" which is "kept separately" and "subject to technical and organisational measures". Hence safety measures such as a secure key management and the respective encryption method used by the controller have to be used "to ensure that the personal data are not attributed to an identified or identifiable natural person". Therefore, because of its existing assignment rule encryption is an example of pseudonymisation. [158]

However, it is controversial whether encrypted personal data, and thus pseudonymised data, can be regarded as anonymised[159] data. Encrypted personal data should nevertheless undisputedly remain personal data to a person who holds the decryption key.[160] The relevant question is whether encrypted data shall also be personal data for a controller or processor who does not have access to the decryption key, for instance a cloud provider. According to Recital 26 S. 2 GDPR pseudonymised data, which could be attributed to a natural person should be considered to be information on an identifiable natural person. At first sight, this is a clear statement of the EU legislator that pseudonymised data shall always be personal data. Nevertheless, to resolve this dispute, once again the question is crucial whether an absolute or a relative approach regarding the identifiability of a data subject has to be applied. According to the absolute approach (see 2.2.1.4), encrypted data will consequently always be personal data, because somebody, at least the key holder or any other party given sufficient time, economic resources and computing power, will always be able to decrypt the data, since no system of encryption can be completely secure[161]. According to this logic, encryption is merely a technical and

[154] *Karg*, DuD 20155, 520 (522).

[155] *Zuiderveen Borgesius*, Computer Law & Security Review 2016, 256 (267).

[156] According to Recital 78 GDPR, personal data should be pseudonymised "as soon as possible".

[157] *Marnau*, DuD 2016, 428 (431).

[158] *Art. 29-Working Party*, Opinion 05/2014, WP 216, 21; *Esayas*, European Journal of Law and Technology, Vol 6, No 2 (2015), 8; *Hennrichs*, Cloud Computing, 2016, p. 137.

[159] For an overview of existing anonymization techniques such as randomization or generalization see the Opinion of the *Art. 29-Working Party*, Opinion 05/2014, WP 216, 12 ff.; *Lagos*, Indiana Law Review 2014-2015, 187 ff.

[160] *Art. 29-Working Party*, Opinion 05/2014, WP 216, 29; *Borges*, in: Borges/Meents (eds.), Cloud Computing, 2016, 6 Recital 33; *Polonetsky/Tene/Finch*, Santa Clara Law Review, (Forthcoming) 2016, 593 (613).

[161] Cf. *Kuner*, International Business Lawyer 1996, 186.

organisational measure to ensure that data is not accessible to unauthorised persons rather than changing the data's quality. However, with a relative approach the data could be regarded as anonymous for the controller.

Consequently, we have to examine which level of encryption is sufficient so that with a relative approach the encrypted personal data can be considered as anonymous data. As mentioned above, only the knowledge and possibilities of the controller to identify the data subject shall be taken into account, therefore, processing encrypted data without affecting the scope of the data protection law might be possible. [162]

In order to concretise whether the means to decrypt the dataset and identify the data subject are reasonably likely to be used and thus whether the encryption method can be regarded as computationally secure or not, one should take account of objective factors. There are three relevant factors that have to be considered when assessing the level of security of encrypted data and against decryption, namely the strength of the encryption algorithm used, the length of the encryption key (the longer the key the safer the encryption will be) and the security of the key management. [163] Obviously, the key always has to be stored separately from the encrypted data in a secure way. If not, attackers may easily be able to decrypt the data[164] and thus, the personal data would no longer be anonymous. The simplest and most common way of decryption is using exhaustive key search or brute-force attacks which means to try all possible keys and eventually guessing correctly. [165] However, if a secure encryption technology is used, this way of decrypting the dataset cannot be considered as very likely for the controller. [166]

Other approaches to get access to the secret key are e.g. legally getting access to a key via a court decision, extracting the key from software or hardware, or by using accidental errors or systematic backdoors implemented in the encryption technique for law enforcement. [167] These ways are only considered to be likely for the controller if they do not violate the law or if they can be achieved by the use of computational power which can be reasonably expected. However, if a *backdoor* is implemented by the government into an encryption technology, the GDPR would be applicable for the controller who knows about this (governmental) possibility of accessing the personal data.

Additionally, the available encryption technology at the time of the processing has to be considered: applying the ECJ's opinion on encryption (see 2.2.1.4.1) it would not be reasonably likely if it were practically impossible to decrypt the dataset, thus, if a state of the art encryption technology is enabled, in most of the cases, decrypting will be virtually impossible and therefore not likely and only possible with unreasonable efforts. [168]

Arguments against a wide interpretation could be sustained by Recital 57 GDPR, which deals with the data subject's right to access personal data held by the controller, where "the personal data processed

[162] Cf. *Hon/Kosta/Millard/Stefanatou*, Tilburg Law School Legal Studies Research Paper Series No. 07/2014, 10; *Borges*, in: Borges/Meents (eds.), Cloud Computing, 2016, 6 Recital 33; *Hennrichs*, Cloud Computing, 2016, 137.

[163] *Hon/Millard/Walden*, Queen Mary University of London - Legal Studies Research Paper No. 75/2011, 22.

[164] *Art. 29-Working Party*, Opinion 05/2014, WP 216, 22; *Hon/Kosta/Millard/Stefanatou*, Tilburg Law School Legal Studies Research Paper Series No. 07/2014, 10.

[165] See with further examples *Gürses/Preenel*, in: van der Sloot/Broeders/Schrijvers (eds.), Exploring the Boundaries of Big Data, 2016, Part I, 3, Cryptology and Privacy in the Context of Big Data, 49 (62); Kroschwald, ZD 2014, 75 (77).

[166] *Cahsor/Sorge*, in: Borges/Meents (eds.), Cloud Computing, 2016, 10 Recital 32, who state that using the 128 bits key lengths of AES encryption would make such an attack nearly impossible and thus not likely.

[167] *Gürses/Preenel*, in: van der Sloot/Broeders/Schrijvers (eds.), Exploring the Boundaries of Big Data, 2016, Part I, 3, Cryptology and Privacy in the Context of Big Data, 49 (63); the German and French government are currently deliberating on legal obligations to implement backdoors in encryption techniques for law enforcement reasons, see `http://www.interieur.gouv.fr/Actualites/L-actu-du-Ministere/Initiative-franco-allemande-sur-la-securite-interieure-en-Europe`.

[168] Different opinion: *Art. 29-Working Party*, Opinion 05/2014, WP 216, 10, according to which the intentions of the data controller or recipient shall not matter, as long as the data are identifiable, data protection rules shall apply.

by a controller do not permit the controller to identify a natural person". Then, "the data controller should not be obliged to acquire additional information in order to identify the data subject for the sole purpose of complying with any provision of this Regulation". This could be a hint against a too wide interpretation of getting access to a key obtained by a third party.

Additionally, future technological developments of decryption, e.g. due to more computing power or improved algorithms have to be considered (cf. Recital 26 GDPR), especially regarding the lengths of the secret key. The controller has to assess whether the future development is evidently foreseeable and thus ought to be regarded as a present information. [169] The controller should take into account the technological development for the period of time in which the data is meant to be processed, therefore, if the data shall be processed for ten years, he or she has to take the technological possibilities for these ten years into account; if the data can be decrypted in the ninth year, the data shall become personal data from that date on only. [170]

Therefore, due to technical developments, encrypted data will only be anonymous for a certain period of time and thus, the level of encryption has to be checked constantly by the controller and not only when the controller processes the data for the first time. [171] Moreover, if a controller receives an already encrypted dataset, he or she has to obtain further information regarding whether the original dataset has included personal data; if yes, the controller has to regularly check the state-of-the-art of the encryption technique. [172] Thus, if the controller does not have the key to decrypt the data or other means to make it legible, it is in most cases reasonably likely that he or she cannot access the personal information, which consequently has to be regarded as anonymous information.

Therefore, according to the GDPR, when using state-of-the-art encryption techniques, encrypted personal data can be anonymous data, with the limitation that a potential possibility of obtaining the key, also by a third party and especially due to decryption, always has to be considered, but only if the means used are reasonably likely.

2.2.6 Summary and Impact on Cloud Computing and Encryption

In sum, the GDPR includes both absolute and relative elements. Consequently, with regards to cloud computing, the GDPR will in many cases not be applicable to the cloud provider if secure encryption technologies are used. For the controller the data nevertheless remains personal data (see in detail 2.3.2), thus this approach leads to a partial non-applicability of the GDPR. However, considering the absolute elements of the Regulation and the decision of the ECJ, supplementary knowledge of third persons should be considered, but only if they are reasonably likely to link the data to an individual. Therefore, encryption of personal data could be a way to anonymize personal data, provided that it is on an adequate level and state-of-the-art and provided that it is not reasonably likely that a third person could re-identify the data to the data subject. The risks of re-identification nevertheless still exist and the controller has to ensure an adequate level of encryption.

[169] *Borges*, in: Borges/Meents (eds.), Cloud Computing, 2016, 6 Recital 38.

[170] *Art. 29-Working Party*, Opinion 05/2014, WP, 18.

[171] *Borges*, in: Borges/Meents (eds.), Cloud Computing, 2016, 6 Recital 40; different opinion Lundevall-Unger/Tranvik, International Journal of Law and Information Technology 2010, 53 (71) who call it "a burden [for the controllers] that they probably cannot be expected to bear" and state that it "will not make controllers in a wired world more inclined to comply with the provisions of the [European data protection law]".

[172] *Borges*, in: Borges/Meents (eds.), Cloud Computing, 2016, 6 Recital 41.

2.3 The Responsible Party (the *Controller*) and Processing on behalf of the Controller

2.3.1 The Responsible Party (the Controller) and Processing on behalf of the Controller under the DPD

2.3.1.1 Relevance

The data protection law addresses the consequence of being a controller. All requirements needed to fulfill compliance with the data protection law have to be ensured by the controller, and possible fees and court rulings will apply to them.
Concerning cloud computing, there can be a lot of entities involved in the whole process of storing and using data in the cloud. For a legal evaluation it is crucial to determine the respective controller. Whereas the cloud user might have clients whose data they are working with, the cloud provider might have sub-contractors whose resources they are using when their own capabilities are limited. [173] One should distinguish between "single" controllers, joint controllers, processors, and third parties.

2.3.1.2 The Controller

Defined as the "natural or legal person that is alone, or jointly with others, responsible for the processing of data," a "data controller" determines the purposes and means of the processing, Article 2 (d) DPD. It is not necessary for the controller themselves to process the data (see 2.3.1.4). It is necessary to describe in detail two important elements included in this definition. First, the controller is the *determining* element - the one who makes the decisions - with respect to the specific data processing action. Second, the subjects left to the controller's determination are the *purposes* and *means* of the processing. The element of determination is a matter mainly based upon factual control, which arises out of the circumstances of the concrete situation. Assessing those circumstances, the controlling-capacity might be derived from explicit legal competence, if one entity is either explicitly appointed as a controller or is imposed with particular data processing duties by legal provisions. It might also be indicated by traditional roles, which usually involve certain data responsibilities, e.g. the collection of specific information about employees by the employer. Finally, the factual influence has to be assessed. For this purpose, the contractual relations between the parties can be analysed. An important indication could be whether the role of the controller is assigned to one party, or whether this party can be considered dominant relating to data issues altogether. However, contractual provisions are not decisive in every case - especially if they do not reflect the factual circumstances. Where doubts occur, the actual control of the parties has to be measured and assessed, taking into consideration the degree of influence actually exercised and the reasonable expectations of the data subjects concerned. [174]

2.3.1.3 Joint Controlling

In a more simplified data processing situation, there might only be one party held responsible when relating to the processing action, such as a controller. Nevertheless, the definition provided within Article 2 (d) expressly includes "control jointly executed by more than one entity." In scenarios where many parties are involved, it is conceivable that various entities can take on the role of joint

[173] *Brennscheidt*, Cloud Computing, p. 59.
[174] *Art. 29-Working Party*, Opinion 01/2010, WP 169, 8 ff.

controllers. As a consequence, each of these parties are bound to the provisions stated within the DPD, with respect to the entire processing action. [175]

The general criteria to assess this form of controlling are, in principle, the same as for "normal" controlling of only one party (see 2.3.1.2). [176] In other words, two or more parties are joint controllers if they determine the essential means and the purposes of the data processing solely together. [177] However, in practice, the line between joint controlling, on the one hand, and order processing (see 2.3.1.4.2) of data, on the other hand, is blurred - and often leads to quarrels with supervisory authorities.

The entities do not need to have a close relationship to each other- for instance, a civil partnership or similar close contractual relations. The parties can generally choose any legal form to establish their relationship - though, this does not affect the responsibility imposed by data protection law. [178] However, contractual agreements can contain important indications for assessing joint controlling (as well as for "single" controlling, see 2.3.1.2) in many cases. Nevertheless, a complete assessment of all specific circumstances is required in order to decide whether parties should take the decisions jointly, or if only one party has to be regarded a ("single") controller. [179] Therefore, it is not important who has the formal right to decide what happens with the data, rather it is crucial who has the actual competence to determine the purposes and means of the processing. [180]

The legal assessment is unambiguous regarding where the different parties jointly determine both the purposes and the means of one particular processing action. However, the Art. 29 Working Party's opinion includes a broader approach to define the scope of joint controlling. According to this opinion, it should be noted, that joint controllers do not need to share the same purposes of the processing - they might differ. Depending on the situation, it either suffices if they only set up an infrastructure of data processing and determine the essential elements of the means to be used or if they share the same purpose without jointly deciding on the means. [181]

Furthermore, as the Art. 29 Working Party argues, the question of joint controlling is not a matter of one particular data processing action. As Article 2 (b) DPD states, the term "processing" is not limited to one single action but also includes a "set of operations" (see 2.1.4). [182]

Especially in the context of IT-infrastructures, there can be many parties involved in different data processing operations of a particular set of personal data. A distinction has to be made if those parties are either "single" controllers that are independent from each other or if they are joint controllers (or if it is a case of order processing, see 2.3.1.4). It is possible that the involved parties divide different tasks and processing operations in such a way so that each single action appears to be independent and executed by only one controller. However, by taking into consideration the whole set of operations - the "macro-level" - the entities can also be regarded as joint controllers. This result can be derived from mutually determined purposes and a cooperatively set framework that determines the essential means or whether the decisions relating to both questions are taken together. [183]

Again, the question of joint controlling is - as with respect to "single" controlling - a matter of the specific circumstances if the parties factually determine the purposes and/or essential means together. Though many different scenarios with different legal assessment can occur, one example may illustrate

[175] *Wolff/Brink*, Datenschutz in Bund und Ländern, par. 3, recital 112.

[176] *Art. 29-Working Party*, Opinion 01/2010, WP 169, 18.

[177] *Art. 29-Working Party*, Opinion 01/2010, WP 169, 18; *Funke/Wittmann*, ZD 2013, 211 f.; see also: *Alich/Nolte*, CR 2011, 741, (743 f).

[178] *Dammann*, in: Simitis, BDSG, par. 3, Recital 226.

[179] *Art. 29-Working Party*, Opinion 01/2010, WP 169, 18; see also 2.3.1.2.

[180] *Jandt/Roßnagel*, ZD 2011, 160, *Jotzo*, MMR 2009, 232 f.

[181] *Art. 29-Working Party*, Opinion 01/2010, WP 169, 19 f.

[182] *Art. 29-Working Party*, Opinion 01/2010, WP 169, 18.

[183] *Art. 29-Working Party*, Opinion 01/2010, WP 169, 20.

the issue: [184] An airline, a hotel chain and a travel agency establish a platform provided through the internet that allows enhanced collaborative travel reservation management between them. They jointly state which data are to be stored on the platform, how reservations are managed and confirmed, to whom access to the data shall be granted, etc. Here, all three parties are joint controllers, with respect to the processing executed by using the common internet-platform, since they decided, at least, about the essential means of the processing.

However, one should bear in mind, that the Art. 29-Working Party opinions have no binding statements (see Article 29 section 1 DPD). In particular, it may be subject to further discussion if such a broad understanding of joint controlling can generally be accepted. The ECJ's recent Google Spain judgment seems to embrace such an understanding. A joint controllership was assumed without the controllers intending to cooperate or jointly deciding on the purpose of the data processing. [185] Simply the fact that both parties were able to control the processing had been sufficient for the ECJ to assume joint controllership. [186]

In a usual cloud computing scenario, the cloud-provider does not determine the means and purposes of the data processing, and there is usually no controller at all (see 2.3.2.4.2). Hence, joint controlling might occur with respect to cases in which more than one user controls the processing action by taking these decisions jointly.

2.3.1.4 Processing on behalf of the Controller

2.3.1.4.1 The Processor

As mentioned above, the controller does not necessarily have to be the entity actually processing the data. On the contrary, companies whose main business is outside the IT-sector tend to outsource data processing. According to the law, a "processor" is any legal entity processing the data on behalf of the controller (Article 2 (d) of the DPD) - the outsourcing company. All data processing the processor does, is considered as processing done by the controller (the outsourcing company) whose responsibility relating to these processing actions is not affected. As a consequence, all given consent and all legal permissions that the controller has are valid to permit the processor's actions regarding personal data. The processor is treated as if they belonged to the controller's entity. Therefore, no permission is needed for data transfers between the controller and the processor. Sometimes this scenario is also called "order processing".

Acting "on behalf" of the controller contains two basic elements: on the one hand, a processor acts in the controller's interests and not for their own purposes. On the other hand, they are bound to the controller's instructions (see Article 16 DPD), at least with respect to the purposes of the processing and the essential means that are used. In this context, the purpose is the "anticipated outcome that is intended or that guides your planned actions" and the means can be defined as "how a result is obtained or an end is achieved". [187] Furthermore, only an entity legally separated from the controller is in general able to act as a processor. [188]

2.3.1.4.2 Distinction between Processor and Controller

Whenever one entity processes (personal) data for another one, the question that arises is whether or not the one actually processing has to be considered a controller or a processor. The distinction

[184] *Art. 29-Working Party*, Opinion 01/2010, WP 169, 20.

[185] *ECJ*, Judgment from the 13th May 2014 in Casec-131/12 - Google Spain SL/Google Inc. v AEPD/Mario Costeja Gonzalez, Recital 40.

[186] Cf. *Spindler*, JZ 2014, 981 (983).

[187] *Art. 29-Working Party*, Opinion 01/2010, WP 169, 13f., 25.

[188] *Art. 29-Working Party*, Opinion 01/2010, WP 169, 25.

between these two roles should be carried out on the basis of the potential control of the party in question. That means, that whoever fulfills the described conditions of being a controller is regarded as a controller and not as a processor (and - of course - neither as a third party). [189] Thus, if one determines the purposes and essential means (at least by giving instructions) he is a controller. [190] In this context, it is crucial to specify which particular decisions can be delegated to the processor, and in contrast, how much leeway or discretion is assigned to the processing party so that it can be considered as a controller rather than a mere processing party, due to the freedom to decide upon specific means of data processing etc.

The possible decisions that are subject to delegation can be divided into two categories requiring different legal assessment: Decisions concerning the purpose of the processing cannot be delegated and are reserved for the controller's authority only. [191] As a consequence, the cloud service provider will be considered a controller if they collect their users' personal data for their own purposes. [192] Decisions that concern the means of the processing, such as which software should be used, may on the other hand be delegated to the processor. However, this does not encompass every technical or organizational question. Some are deeply linked to the lawfulness of the processing and, therefore, essential in a way that they can only be answered by the controller. In particular, this is especially relevant to aspects such as the duration of the processing, granting access to third persons, and the choice of which data should be processed. [193]

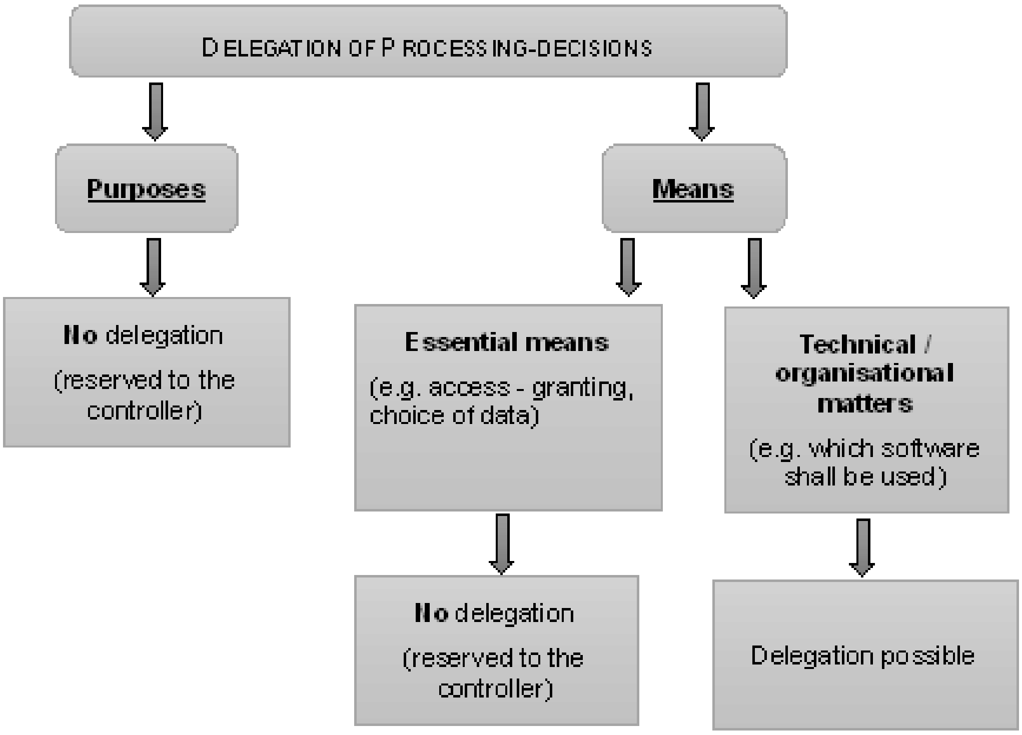

DELEGATION OF PROCESSING-DECISIONS
Purposes
Means
No delegation
(reserved to the controller)
Essential means
(e.g. access - granting, choice of data)
Technical / organisational matters
(e.g. which software shall be used)
No delegation
(reserved to the controller)
Delegation possible

Figure 2.4: Order-processing - Delegation of decision

In a typical cloud computing scenario, the provider only supplies the controller with the technical

[189] *Brennscheidt*, Cloud Computing und Datenschutz, p. 67; *Gola/Klug/Körffer*, in: Gola/Schomerus, Bundesdatenschutzgesetz, par. 11, recital 9.

[190] Cf. *Hilber*, Handbuch Cloud Computing, p. 350.

[191] *Art. 29-Working Party*, Opinion 01/2010, WP 169, 15 f.

[192] *Giedke*, Cloud Computing, p. 202; *Art. 29-Working Party*, Opinion 08/2010, WP 179, 27; *Art. 29-Working Party*, Opinion 05/2012 WP 196, 10.

[193] *Art. 29-Working Party*, Opinion 01/2010, WP 169, 14.

framework. The latter is the one determining the purposes of the processing. Usually, the controller decides which data are processed and how long the processing will take and, therefore, governs the (essential) means. However, the cloud provider only computes the data, as they are bound by the contract concluded with the cloud user, thus possessing little discretionary power that, normally, does not lead to a controllership. [194]

Even though cloud computing can, therefore, typically be regarded as processing on behalf of the controller, in terms of Article 16 DPD, [195] it is deliberated whether or not there can be scenarios in which the provider acts neither as a processor nor as a controller. It is possible that the cloud user does not give any instructions to the cloud service provider on how to handle the data. One might only use the provider's software in a SaaS solution to compute over self-processed input and receive the results. The provider does not exercise any data processing, but only establishes and maintains the technology to support data processing that is completely initiated and conducted by the controller themselves. In such cases, it is argued that one does not "process" on behalf of another but is only indirectly concerned with the data processing and, thus, cannot be considered a processor. [196] Others argue that, under those circumstances, the provisions for data processors apply as well, since the risks for the personal data do not differ significantly when compared to a situation in which the processor directly processes the data. [197] At the very least, the provider's mere physical control over the data requires the implementation of sufficient safeguards to sustain data security in those cases (assumed one shares that approach), for instance measures to prevent data from accidental loss. [198]

However, this discussion should not be overstated. It is important to bear in mind that whenever a cloud-service includes any form of data storage (on the provider's servers), which goes beyond a mere temporary caching, then this storage constitutes a relevant act of data processing. Accordingly, the provider has to be considered a processor. [199] This applies to an even greater extent if the provider fulfills monitoring tasks with respect to the personal data, e.g. concerning the access or use. [200]

However, there can be situations in which the provider fulfills the requirements of controlling and therefore acts as a controller, and not as a processor. A few examples shall be emphasized. In one instance, a former processor starts processing data for their own purposes, or others', other than those originally determined by the (former) controller. For example, if the "processor" starts to use stored customer data in order to provide commercial advertising in a manner not intended by the user, with respect to this new processing action, they are a controller, since they set a new purpose. [201] The same might apply if they exceed other competences, such as granting data access to unauthorized third parties. [202] Furthermore, the provider could be assigned not only with providing the technical framework but also with completing the whole task that leads to the processing action. Whenever the

[194] *Brennscheidt*, Cloud Computing, p. 67 f.; *Hennrich*, CR 2011, 546 (548); cf. also *Wolff/Brink*, Datenschutz in Bund und Ländern, par. 3 BDSG, recital 111; *Niemann/Paul*, Praxishandbuch Rechtsfragen des Cloud Computing, chapter D, recital 31 ff.

[195] *Brennscheidt*, Cloud Computing, p. 67 f; apparently assumed in: *Art. 29-Working Party*, Opinion 05/2012, WP 196.

[196] *Hon/Millard/Walden*, Who is Responsible for "personal data" in Cloud Computing? - The Cloud of Unknowing, Part 2, p. 17; *Spindler/Nink*, in: Spindler/Schuster, Recht der elektronischen Medien, par. 11 BDSG, recital 8.

[197] Cf.*Schneider*, Handbuch des EDV-Rechts, chapter B, recital 266 f.

[198] *Hon/Millard/Walden*, Who is Responsible for "personal data" in Cloud Computing?, The Cloud of Unknowing, Part 2, p. 22.

[199] *Pohle/Ammann*, K&R 2009, 625 (630); *Spindler/Nink*, in: Spindler/Schuster, Recht der elektronischen Medien, par. 11 BDSG, recital 8; see also more differentiated if the provider has a mere passive role: *Hon/Millard/Walden*, Who is Responsible for "personal data" in Cloud Computing? The Cloud of Unknowing, Part 2, p. 18 ff.

[200] *Hon/Millard/Walden*, Who is Responsible for "personal data" in Cloud Computing?, The Cloud of Unknowing, Part 2, p. 17.

[201] *Art. 29-Working Party*, Opinion 05/2012, WP 196, 14.

[202] *Hon/Millard/Walden*, Who is Responsible for "personal data" in Cloud Computing?, The Cloud of Unknowing, Part 2, p. 20.

provider is empowered with the competences to decide the essential means and purposes with respect to that task, they are a controller - even though the involved parties may consider them rather as a processor. [203] The outsourcing of a company's accountancy is a typical example, for this respect. [204]

2.3.1.4.3 Legal Requirements

There are certain legal requirements to fulfill before (order) processing takes place "on behalf of the controller" (Article 17 Par. 3 DPD); for example, the carrying out of the processing must be governed by a contract or legal act binding the processor to the controller. The processor must be bound to instructions from the controller, and it must be guaranteed that technical and organizational measures are provided to protect personal data against leaks. The main aim is to obligate the processor to follow the controller's instructions, similar to an employee's obligation. For the purposes of verification, the sections of the contract or the legal act relating to data protection and the requirements concerning the technical and organizational measures shall be in writing or another equivalent form. [205] One may note that users, especially small cloud users, usually do not have a considerable influence on the contractual clauses often provided in a standardized form by the provider. However, it remains part of the controller's responsibility to only enter into processing-contracts which are in complete compliance with the respective legal data protection provision. A lack of actual power does not justify concluding an unlawful processing-contract. [206]
The EU's Art. 29-Working Party recommends certain issues to be covered in a contract between the cloud provider and the user. For example they provide:

- for details concerning the client's instructions to be issued to the provider and
- relevant penalties, including potential actions against the provider, in case of non-compliance,
- specification of the security measures the provider must comply with,
- subject and time frame of the cloud service to be provided,
- a confidentiality clause,
- the controller's rights to monitor,
- the cloud provider's obligation to cooperate,
- a list of locations in which the data may be processed, and
- the prohibition of communicating data to third parties or subcontractors not mentioned in the contract. [207]

From a critical point of view, these requirements are difficult to fulfill in practice. On the one hand, it is highly unlikely that big global players in the cloud computing business will actually be bound and controlled by mid-sized or small companies concerning cloud computation (for instance, referring to

[203] Cf.*Brennscheidt*, Cloud Computing und Datenschutz, p. 67; *Funke/Wittmann*, ZD 2013, 221 (223); *Hoeren*, in: Roßnagel, Handbuch Datenschutzrecht, chapter 4.6, recital 97.

[204] *Petri*, in: Simitis, BDSG, par. 11, recital 28.

[205] *Art. 29-Working Party*, Opinion 05/2012, WP 196, 12.

[206] *Art. 29-Working Party*, Opinion 01/2010, WP 169, 26; Hartung/Storm, in: Hilber, Handbuch Cloud Computing, p. 357.

[207] The whole list of recommendations has 14 items and can be found in Art. 29-Working Party, Opinion 05/2012, WP 196, 12 f.

inspections on the spot). On the other hand, a company not operating in the IT-sector might not even be interested in or be able to provide this kind of control. [208] Since the data might be stored not in one but in many different locations, visiting the provider's data centres for an on-site audit seems implausible for the cloud user. In addition, it might even be hard to tell where exactly the data will be stored due to the scalability of cloud services. [209] Besides the difficulties for a cloud user to visit and audit all data centres his provider is using, it would constitute a data-security risk for the provider to let (all of) their users inspect all their data centres. This model of control is based upon the classic outsourcing model, with only one data centre to be controlled that might not be located in another country. However, other options to fulfill the cloud user's legal obligations to control his provider have been proposed: As the directive does not require the controller to ensure the processor's compliance by themselves, they could rely on a qualified third party to control the processor (third-party auditing model). [210] On the other hand, the cloud user would still have to pay for the services of this third party, something that might be impractical even for private individuals. The controller could demand inspection reports from the processor recording his processing activities, but this would not ensure the processor's actual compliance, since those reports would be made by the processor themselves. [211] An effective, yet practical, way to ensure compliance is data protection certification. [212] Here, a third party provides the necessary assessment of the cloud provider. Compared to the third party audit-model mentioned before, the difference is that not every client of the provider has to hire the third party individually. The certification costs are initially covered by the cloud provider and then redistributed to all possible clients by the provider - making it possible to professionally control every data centre, as well as affordable for private customers. Being certified might provide a competitive advantage for big, global players since this advertises a high standard of data protection to possible clients. The directive does not mention such certificates explicitly, nevertheless, they could be used by a controller to ensure the compliance of the processing done on their behalf. [213]

2.3.1.4.4 By Processor Outside the EU/ EEA

If the processor does not fall under the jurisdiction of an EU/EEA member-state, data transmission between the controller and the processor generally have to comply with the described conditions. In addition, the requirements of data transfer to third countries have to be met (for more details see 2.4.3); under no circumstances shall personal data be transferred to a third country that is not providing an adequate level of protection without the described requirements (see 2.4.3). Nevertheless, the contract binding of the processor to the controller can be used to ensure necessary safeguards. Thus in sum, only if either an adequate level of protection is provided within the third country or other sufficient safeguards are ensured, will the DPD allow it to constitute an order processing, including the legal privileges described in 2.3.1.4.1.[214]

[208] *Heidrich/Wegener*, MMR 2010,803 (806).

[209] *Brennscheidt*, Cloud Computing, p. 102.

[210] *German Federal Office for Information Security Technology*, Safety Recommendation for Cloud Computing Providers, p. 63

[211] *Brennscheidt*, Cloud Computing, p. 105.

[212] *Art. 29-Working Party*, Opinion 05/2012, WP 196, 22.

[213] For a detailed description of data protection seals see *Brennscheidt*, Cloud Computing, p. 105 ff.

[214] *Brennscheidt*, Cloud Computing, p. 76.

2.3.2 The Responsible Party (the Controller) and Processing on behalf of the Controller under the GDPR

In the same way, the GDPR distinguishes between the entity responsible and the entity actually processing the data. Nevertheless, there are changes in the particular responsibilities of those entities and new ways for the controller to make sure his or her processor complies with the law. It is essential that order processing under the GDPR meets all prerequisites described below.

It has been criticized that there is no regulation within the GDPR that explicitly states that transfers from a controller to the processor are allowed if "order processing" takes place. [215] Yet, this critique does not take into account that the legitimation for such transfers lies in the model of "order processing," itself. Without this legitimation, all provisions regarding processing on behalf of the controller would be meaningless. [216]

In general, the processing of data by the processor (such as computation of cloud-stored data in the cloud) is permitted if the controller would be allowed to do it himself - be it by consent or be it other explicit legal permissions. Hence, data processing is permitted under same circumstances and requirements as for the controller. The processing is done on behalf of the controller, i.e. the law treats the processing as if the controller would do it, himself. Therefore, the controller needs to be the party deciding why and how the processing is done (see 2.3.2.1 ff.).

2.3.2.1 Rules for the Controller

The controller is defined in Article 4 Par. 5 GDPR as

> "the natural or legal person, public authority, agency or any other body which alone, or jointly with others, determines the purposes, conditions and means of the processing of personal data; [. . .]."

In comparison to the DPD, there will be no significant changes to the definition of "controller", for cloud computing. The cloud user as the entity determining the purpose and the means of the data processing will still be considered the controller (for several controllers see 2.3.2.2). The user (= the controller) is thus responsible for the data processing and will be accountable if legal requirements are not met. The controller's main duties are regulated in Article 24 of the GDPR.

> Article 24 Par. 1 S. 1: "Taking 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, the controller shall implement appropriate technical and organisational measures to ensure and to be able to demonstrate that processing is performed in accordance with this Regulation."

Simplified, this means that it is the controller's (the cloud user's) job to ensure that the GDPR's requirements are fulfilled when they initiate data processing. To reach that goal, the controller has to implement technical and organizational measures (see 2.4.4.2) and adopt appropriate policies. To determine if these measures are valid to ensure compliance with the data protection law and the data subjects' privacy, Article 22 Par. 1 provides certain criteria (at least in the LIBE-Proposal).

[215] *Nebel/Richter*, ZD 2012, 407 (411); *Roßnagel/Nebel/Richter*, ZD 2013, 103 (105); c.f. *Koòs/Englisch*, ZD 2014, 276 (284), who see the legitimation in Article 6 lit. f GDPR, if data transfers between the controller and the processor will be considered as necessary for the purposes of the legitimate interests pursued by the controller and not overridden by the interests of the data subject (see 2.4.2.2) and therefore be based on a express legal permission.

[216] C.f. regarding the DPD, but with the same problem: *Drews/Montreal*, PinG 2014, 143.

Using privacy preserving cloud computing technologies developed by PRACTICE can be an efficient way for the controller to comply with these technical and organisational obligations. (see in detail 2.4.4.2)
Aside from the obligation to ensure compliance, the controller also has to be able to demonstrate the adequacy and effectiveness of those measures and policies. To achieve this, Recital 74 stipulates that "those measures should take into account the nature, scope, context and purposes of the processing and the risk for the rights and freedoms of natural persons" and provides in Recital 75 several examples of such risks and sets in Recital 76 that risks "should be evaluated on the basis of an objective assessment, by which it is established whether data processing operations involve a risk or a high risk". The controller may comply with the obligations to implement appropriate measures by using approved codes of conduct or approved certifications (Recital 77, see in detail 2.3.2.4). Article 30 requires documentation (or in the words of the GDPR: "a record of processing activities") of the data processing under the responsibility of the controller. The controller shall according to Article 31, on request, cooperate with the supervisory authority; take technical and organizational measures to ensure the security of processing, Article 32; the controller shall alert and inform clients about data breaches, according to Article 343 Par. 1; conduct a privacy impact assessment under certain conditions of Articles 35, or seek a prior authorization in accordance with Article 36 Par. 1; appoint a data protection officer, as requested in Article 37 Par. 1; as well as comply with rules for transfers to third countries, as mentioned in Article 44 ff. The powers of regulators may be expressly addressed to the processors according to Article 58 Par 1 (a).

2.3.2.2 Joint Controllers

The GDPR's definition of 'controller' allows several entities to be considered as 'joint controllers'. Since the GDPRs definition of 'controller' has only been slightly changed in comparison to the DPDs definition, the distinction between one 'controller' or several 'joint controllers' is still the same (see 2.3.1.4.2) under the DPD. In the case of a 'joint controllers' scenario, it might be difficult to determine the specific responsibilities of each controller. Article 24 GDPR binds joint controllers to come to an arrangement that clarifies each controllers' duties. According to Recital 62 GDPR, the arrangement should reflect the controllers' roles and relationships. The essence of the arrangement has to be made available to the data subject. This is important, since it is necessary that the arrangement determines which controller is responsible for the procedures and mechanisms involved in exercising the rights of the data subject. The reason behind this is that the joint controllers might not be equally capable of negotiating a contract. Additionally, one controller could have a direct relationship to the data subject, whereas another one might not; moreover, they may not be able to control the type and amount of data. [217] The arrangement the GDPR demands should be viewed as a useful tool for cloud participants when they are considered joint controllers. Determining the cloud service's details between the cloud provider and the cloud user(s) may be included in the contract. If the respective responsibilities are not clear to the data subject, all joint controllers are liable, together or separately. In this specific case, the rationale is to provide the data subject with more protection. [218]

[217] *Kelly*, ITRE Committee Opinion on the proposal for a regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation), COM (2012) 0011-C7- 0025/2012-2012/0011 (COD) of 26/02/2013, 102, available at: http://www.europarl.europa.eu/meetdocs/2009_2014/documents/itre/ad/927/927816/927816en.pdf.

[218] *Comi*, IMCO Committee Opinion on the proposal for a regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation), COM (2012) 0011 -C7-0025/2012-2012/0011 (COD) of 28/01/2013, 79, available at: http://www.europarl.europa.eu/meetdocs/2009_2014/documents/imco/ad/924/924645/924645en.pdf.

This underlines Article 24 Par. 2 of the proposal of the Council, which states that "irrespective of the terms of the arrangement the data subject may exercise his or her rights (. . .) in respect of and against each of the controllers".

2.3.2.3 Rules regarding the Processor

In the relationship between the cloud provider and the cloud user, the cloud provider usually acts as the processor as defined in Article 4 Par. 6 (also often called "order processing").

> "(6): 'processor' means a natural or legal person, public authority, agency or any other body which processes personal data on behalf of the controller;"

The cloud user remains accountable to the person (his client) concerned; [219] one should bear in mind that the cloud user often offers services to their clients, thus has to be qualified as a controller (Client - cloud user - cloud provider). The controller's duties regarding the processor commence before the processing on their behalf takes place, and in addition they have to choose a processor who will comply with the GDPR's requirements:

> "Article 28 Par. 1: Where processing is to be carried out on behalf of a controller, the controller shall choose a processor providing sufficient guarantees to implement appropriate technical and organizational measures and procedures in such a way that the processing will meet the requirements of this Regulation and ensure the protection of the rights of the data subject" (emphasis added)

Article 32 GDPR clarifies what is meant by 'technical and organisational measures' (see 2.4.4.2). When assessing the appropriate level of security "account shall be taken in particular of the risks that are presented by processing, in particular from accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to personal data transmitted, stored or otherwise processed" (Par. 2). (Article 30 Par. 2 lit. b). This is possible with encryption technology
If the cloud provider uses privacy-preserving technologies like encryption, such as those developed by PRACTICE, it can be assumed the provider (at least partially) fulfills its duties regarding aforementioned technical security measures, e.g. with the help of certifications or privacy seals (see 2.3.2.4 and also 2.4.4.2 regarding technical and organizational measures).
Of course, the cloud providers still have to take organizational measures to fulfill all of their duties regulated in Article 32 GDPR.
The controller not only has to choose a sufficient processor. The Regulation sticks to the former approach of the DPD and requires the controller to ensure they have control over the data processing (determining the means of the processing, the required organisational and technical measures, processing only on their instructions, their inspection rights, etc.) by contractual obligations of the processor; Article 28 Par. 2 defines a set of rules that must, in practice, be endorsed in the contract or in the other legal act: Article 28 Par. 3 S. 2 GDPR: "That contract or other legal act shall stipulate, in particular, that the processor:

> (a)processes the personal data only on documented instructions from the controller, including with regard to transfers of personal data to a third country or an international organisation, unless required to do so by Union or Member State law to which the processor is subject; in such a case, the processor shall inform the controller of that legal requirement before processing, unless that law prohibits such information on important grounds of public interest;

[219] *Kroschwald*, ZD 2014, 75 (78); *Weichert*, DuD 2010, 679 (682).

> (b) ensures that persons authorised to process the personal data have committed themselves to confidentiality or are under an appropriate statutory obligation of confidentiality;
>
> (c) takes all measures required pursuant to Article 32;
>
> (d) respects the conditions referred to in paragraphs 2 and 4 for engaging another processor;
>
> (e) taking into account the nature of the processing, assists the controller by appropriate technical and organisational measures, insofar as this is possible, for the fulfilment of the controller's obligation to respond to requests for exercising the data subject's rights laid down in Chapter III;
>
> (f) assists the controller in ensuring compliance with the obligations pursuant to Articles 32 to 36 taking into account the nature of processing and the information available to the processor;
>
> (g) at the choice of the controller, deletes or returns all the personal data to the controller after the end of the provision of services relating to processing, and deletes existing copies unless Union or Member State law requires storage of the personal data;
>
> (h) makes available to the controller all information necessary to demonstrate compliance with the obligations laid down in this Article and allow for and contribute to audits, including inspections, conducted by the controller or another auditor mandated by the controller.

With regard to point (h) of the first subparagraph, the processor shall immediately inform the controller if, in its opinion, an instruction infringes this Regulation or other Union or Member State data protection provisions.
Although the data protection law will be renewed, the practical problems will still be the same. The cloud user, or/and the controller, might not be in the position to determine contractual clauses but might have to agree to whatever the much stronger processor (the cloud provider) dictates (see 2.3.1.2.1). It may also be impossible for the cloud user to do on-site inspections for the reasons described above. This problem has been addressed by the GDPR, since it is now possible for the controller to rely on data protection seals and third party audits (see 2.3.2.4).
In contrast to the DPD, the legal consequence of a breach of this agreement is explicitly regulated. Thus, Article 28 Par. 10 states:

> "(...) if a processor processes personal data other than as instructed by the controller or if they become the determining party in relation to the purposes and means of data processing, the processor shall be considered to be a controller in respect of that processing (emphasis added)."

The shift of the processor's role (from mere processing to determining and controlling any data processing) thus leads to a re-qualification of the processor now as a data controller - with all obligations and duties.
Important for cloud computing services is the allowance stated in Article 26 Par. 2 (d) that the processor may use the services of other processors. [220]
According to Article 28 Par. 2 "The processor shall not engage another processor without prior specific or general written authorisation of the controller. In the case of general written authorisation, the

[220] *Brennscheidt*, Cloud Computing, p. 116.

processor shall inform the controller of any intended changes concerning the addition or replacement of other processors, thereby giving the controller the opportunity to object to such changes." Thus, a cloud provider may mandate other sub-contractors (sub-cloud providers, etc.) to process the data. However, the data controller is still in charge of controlling the whole process, so that he or she has to assure that his or her inspection rights, etc., are also enforceable in the relationship with the third-party processor (sub-cloud provider).

Article 28 Par. 4 stipulates these principles as follows: "Where a processor engages another processor for carrying out specific processing activities on behalf of the controller, the same data protection obligations as set out in the contract or other legal act between the controller and the processor (. . .) shall be imposed on that other processor by way of a contract or other legal act under Union or Member State law, in particular providing sufficient guarantees to implement appropriate technical and organisational measures (. . .).Where that other processor fails to fulfill its data protection obligations, the initial processor shall remain fully liable to the controller for the performance of that other processor's obligations." According to Par. 9 the contract or the other legal act referred to in Par. 3 and 4 shall be in writing which also includes the electronic form. This other legal act referred to in Par. 3 and 4 may be based on standard contractual clauses (see Par. 6-8) which may be adopted by a supervisory authority or laid down by the Commission including when they are part of a certification granted to the controller or processor.

Article 29 GDPR regulates that "(t)he processor and any person acting under the authority of the controller or of the processor, who has access to personal data, shall not process those data except on instructions from the controller, unless required to do so by Union or Member State law". Thus, the processor shall act only on the instructions of the controller. However, as outlined above (2.3.2), there is no provision in the GDPR that deals with an explicit exemption of the principle of prohibition with the reservation of permission for the controller and for the processor in an order processing scenario. [221] A solution could be to consider the processing of personal data on behalf of the controller by the processor as a single process of data processing which consequently means that the order processing itself would not need to comply with the regulations of Article 6 Par. 1 GDPR. [222] Thus, if e.g. the storing of personal data by a controller is lawful according to Article 6 Par. 1 GDPR it will also be lawful for the processor to store the data according to Articles 28 and 29 GDPR. Additional obligations of the GDPR regarding processors are:

- designating a representative in the Union according to Article 27 Par. 1 GDPR
- maintaining a record of all categories of processing activities carried out on behalf of a controller according to Article 30 Par. 2 GDPR
- cooperating with the supervisory authority according to Article 31 GDPR
- designating a data protection officer according to Article 37 GDPR
- complying with the principles for personal data transfer to third countries according to Articles 44 ff. GDPR

[221] *Härting*, ITRB 2016, 137 (138). Another possibility is to consider the order processing as an "legitimate interest" of the controller according to Article 6 Par. 1 S. 1 lit. (f) GDPR, however, this could be impractical, since according to Article 21 Par. 1 GDPR the data subject shall have the right to object to a processing based on "legitimate interests" at any time

[222] *Härting*, ITRB 2016, 137 (138 f.)

2.3.2.4 Privacy Seals and Certification

According to Article 42 Par. 1 GDPR, the Member States, the supervisory authorities, the Board and the Commission shall encourage the establishment of data protection certification mechanisms and of data protection seals and marks "(i)n order to enhance transparency and compliance with this Regulation" (Recital 100 GDPR). Each controller and data processor has the right to apply for a certification procedure as mentioned in Article 42 Par. 1 GDPR.

The certification shall be voluntary and available via a process that is transparent. The certification procedure, however, may turn out to be in practice, one of the most important tools for data controllers to present evidence required by Article 28 Par. 1, concerning the selection of processors with sufficient guarantees for data protection, particularly appropriate technical and organisational measures. [223] This could potentially be a solution for the dilemma arising from the disparity of power between the cloud computing participants as the cloud provider will be able to request a certification that the cloud user can rely on. Article 28 Par. 5 GDPR emphasises this by stipulating that "(a)dherence of a processor to an approved code of conduct (...) or an approved certification mechanism (...) may be used as an element by which to demonstrate sufficient guarantees as referred to in paragraphs 1 and 4 of this Article. The certification shall be available via a process that is transparent (see Article 42 Par. 3).

However, Article 42 Par. 4 affirms that:

> "A certification pursuant to this Article does not reduce the responsibility of the controller or the processor for compliance with this Regulation and is without prejudice to the tasks and powers of the supervisory authorities which are competent pursuant to Article 55 or 56." (emphasis added)

Nevertheless, the GDPR settles in Article 42 Par. 1 that certification mechanisms shall be established for the "(...)purpose of demonstrating compliance with this Regulation of processing operations by controllers and processors." (emphasis added) Moreover, administrative authorities can be bound by their previous practices and decisions regarding certifications - however, this binding effect does not apply for those provisions in the GDPR that establish certifications only as a factor of compliance (such as Recital 81 GDPR: "(...) an approved certification mechanism may be used as an element to demonstrate compliance with the obligations of the controller." (emphasis added) or in Articles 24 Par. 3, 25 Par. 3, 28 Par. 5 and 32 Par. 3 GDPR). [224] A certificate of the processor may thus be considered sufficient evidence to verify compliance with the GDPR. [225] Nevertheless, the extent to which certifications have a relieving effect for cloud computing remains unclear. [226] Not only can the processor request a certification, but the controller might have an interest in getting certified as well. A cloud user (as the controller) may be able to prove to his clients that he uses a cloud service that is compliant with data protection law that especially provides sufficient technical and organizational safeguards.

Although the certification of the processor can make it more straightforward for the controller to present evidence required by Article 28 Par. 1 GDPR, a certificate will expire after three years.

> Article 42 Par. 7: "Certification shall be issued to a controller or processor for a maximum period of three years and may be renewed, under the same conditions, provided that the relevant requirements continue to be met. Certification shall be withdrawn (...) where the requirements for the certification are not or are no longer met."

[223] *Brennscheidt*, Cloud Computing, p. 116

[224] *Spindler*, ZD 2016, 407 (409, 412).

[225] *Petri*, ZD 2015, 305 (308).

[226] *Brennscheidt*, Cloud Computing und Datenschutz, p. 116; *Heckmann*, in: jurisPK-Internetrecht, Recital 695.

Prior to relying on a certificate, the processor will therefore at least be obliged to validate if it has expired or not. [227] Nevertheless, the problem that arises from the fact that the cloud user has to ensure by contract that he has control over the provider if an "order processing" shall take place is not solved by a certification. [228] A cloud provider (even if they are a big global player) ought to make it easier for their client - the cloud user - to fulfill their obligations by providing standard contracts for their cloud services that involve the requirements of Article 28 Par. 3 GDPR. This way, a lawful use of the cloud service would be possible for the cloud user, as a controller, if the user wants to compute personal data in the cloud. If lawful usage of a cloud service for the computation of personal data is not possible due to the cloud provider denying a contract with their clients (the cloud users, controllers) that, in a way, binds the cloud provider compliant to Article 28 Par. 3 GDPR due to the provider's more powerful position. This in turn might lead to a disadvantage on the European market once the GDPR comes into effect. Moreover, according to Article 42 Par. 2 GDPR data protection certifications may be a way to lawfully transfer personal data to a third country if controllers or processors are not subject to this Regulation pursuant to Article 3. [229]
Finally, the new Article 43 GDPR introduces several procedures and requirements for certifications and the certification bodies.

2.3.2.5 Liability for Controllers and Processors

According to Art. 82 GDPR, if data has been processed unlawfully, the data subject has the right to claim compensation, even for non-pecuniary damages. Unlike the DPD, it is not only the controller who is liable for such damages. If an 'order processing' takes place, the processor faces liability, too:

> "Article 82 Par 1: Any person who has suffered damage, including non-pecuniary damage, as a result of an unlawful processing operation or of an action incompatible with this Regulation shall have the right to claim compensation from the controller or the processor for the damage suffered."

This could potentially have a huge impact on cloud providers (usually processors on behalf of the controller, the cloud user), as it could be more promising to hold the solvent provider liable for the affected person (usually the cloud user's client) than holding the cloud user liable. The GDPR incorporates the possibility to avoid liability for damages.

> "Article 82 Par 3: The controller or the processor may be exempted from this liability, in whole or in part, if the controller or the processor proves that they are not responsible for the event giving rise to the damage."

Since it is the processor or the controller that needs to prove that they are not responsible for the damage, according to Art. 28, they are both obliged to take the technical and organizational measures that the GDPR demands and fulfill their duty to document the processing. This works to the advantage of the affected person claiming compensation for damages as it is up to the responsibility of the processing parties to prove that they are not liable. On the other hand, the affected person still has to provide evidence for the causation of the unlawful processing for the damages. It has been criticized that this might not be possible for the affected person because he will not have insight into, or be able to document, the controller's or the processor's internal procedures. [230] If several processors or joint controllers caused the damages, this would serve as a further advantage for the affected person as illustrated below:

[227] *Sydow/Kring*, ZD 2014, 271 (275).
[228] *Sydow/Kring*, ZD 2014, 271 (275).
[229] See for more details 2.4.3.2 and *Spindler*, ZD 2016, 407 (410).
[230] *Roßnagel/Richter/Nebel*, ZD 2013, 103 (108).

> “ Article 82 Par. 4: “Where more than one controller or processor, or both a controller and a processor, are involved in the same processing and where they are, under paragraphs 2 and 3, responsible for any damage caused by processing, each controller or processor shall be held liable for the entire damage in order to ensure effective compensation of the data subject”

The possibility of joint liability for joint controllers demonstrates the importance of both parties coming to an agreement that fully reflects their responsibilities in the context of data processing. This way, only the respective controller will be liable for possible damages caused by his actions.
However, the GDPR grants a privilege to processors who are not responsible for the damage caused by the processing of a controller:

> Article 82 Par. 2: “ Any controller involved in processing shall be liable for the damage caused by processing which infringes this Regulation. A processor shall be liable for the damage caused by processing only where it has not complied with obligations of this Regulation specifically directed to processors or where it has acted outside or contrary to lawful instructions of the controller.”

This exception above is a positive provision for cloud computing providers who act as processors. The processor shall be exempted from liability if they are able to prove that they are not in any way responsible for the damage (Par. 3, see above). However, the processor will only be exempted from liability if the instructions of the controller have been lawful; if, in the processor’s opinion, an instruction infringes the GDPR or other Union or Member State data protection provisions and he or she would according to Article 28 Par. 3 GDPR be obliged to immediately inform the controller about this issue, the processor will be liable and not be able to refer to the instruction given by the controller. [231]

In contrast to the strict regulations of the LIBE-Proposal, the exception above is a positive provision for cloud computing providers who act as processors. The processor shall be exempted from liability if they are able to prove that they are not in any way responsible for the damage (Par. 3).
If a controller or processor is liable for the damage, they can claim back parts of the compensation from the other responsible party in accordance to par. 5:

> “Where a controller or processor has (. . .) paid full compensation for the damage suffered, that controller or processor shall be entitled to claim back from the other controllers or processors involved in the same processing that part of the compensation corresponding to their part of responsibility for the damage in accordance with the conditions set out in paragraph 2.”

2.4 Requirements for Legal Data Processing

2.4.1 The Definition of ‘Processing’

The Data Protection Directive defines the “processing of data” as:

> “any operation or set of operations which is performed upon personal data, whether or not by automatic means, such as collection, recording, organization, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction’, Art. 2 (b). ”

[231] *Becker*, in: Plath, BDSG, 2nd Ed. 2016, Art. 82 DSGVO, Recital 6.

The exceedingly broad definition of ‘processing’ leads to the applicability of the DPD and, in turn, to the general prohibition of processing the data unless the DPD allows for it. From the moment the data is collected by the data subject to the very last use of that data, every single step in between has to be either explicitly allowed by law or needs the data subject’s consent.

Thus, data controllers can only avoid the applicability of the DPD by rendering the data “not personal”. Otherwise, they can have a duty to comply with the requirements - asking the user for explicit consent or presenting reasons that fall under the justifications provided by the DPD.

If personal data is anonymized, this may, technically, mean that it gets altered, but, for the purposes of the Data Protection Directive, ’alteration’ means changing the content of the information, not its appearance. [232]

Since the anonymization of personal data eliminates the connection to a person, the encryption of data is one of few possibilities to anonymize personal data and, therefore, the process of encrypting data does not fall under the data protection law, either.

The GDPR defines processing in Article 4 No 2 as:

> “(...) any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction”. (emphasis added)

The GDPR’s definition does not include major changes to the former definition of the DPD, it only adds “structuring” of personal data as an example for processing and changes “blocking“ into “restriction”. Thus, the basic principles of the DPD regarding the definition of “processing” will also apply for the GDPR.

2.4.2 Informed Consent or Explicit Legal Permission

2.4.2.1 Legal Permissions in the DPD

Article 7 DPD enumerates the possible legal grounds for data processing. The first possibility is the affected person’s informed consent (see 2.4.2.3). If the controller has not gained the data subject’s consent, a lawful processing is possible on the grounds of one of the legal permissions stated in Article 7 (b) to (f) DPD:

> “[...]
>
> (b) processing is necessary for the performance of a contract to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract; or
>
> (c) processing is necessary for compliance with a legal obligation to which the controller is subject; or
>
> (d) processing is necessary in order to protect the vital interests of the data subject; or
>
> (e) 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 or in a third party to whom the data are disclosed; or
>
> (f) processing is necessary for the purposes of the legitimate interests pursued by the controller or by the third party, or parties, to whom the data are disclosed, except where

[232] See also *Gola/Klug/Körffer*, in: Gola/Schomerus, Bundesdatenschutzgesetz, par. 3, recital 30 f.

> such interests are overridden by the interests for fundamental rights and freedoms of the data subject, which require protection under Article 1 (1)."

Those permissions transferred by the Member States into national law are conclusive, meaning that they are not simply examples among other possible legal grounds but also the only lawful means to process data without the data subjects consent. They permit processing only when it is necessary for certain purposes and not beyond that, corresponding with the DPDs fundamental principle of proportionality laid down in Article 6 DPD (see 2.1.1.2.2.3). Whereas lit (b) to (e) are applicable only to specific purposes, lit (f) allows the Member States to provide a legal grounding with a larger scope. Nonetheless, heeding Article 6 DPDs main principles, processing on the grounds of a permission based on lit (f) always requires a proportionality test. This indicates that balance has to be struck between the data subjects' and the controllers' interests. Only when the controllers' interests in processing the data without consent outweigh the data subjects' interests in having to consent to the processing, can the processing can be lawfully done on the grounds of lit (f). Since gaining informed consent can be difficult for data processing with cloud computing (see 2.4.2.3), it would theoretically be useful if data could be processed without consent. Cloud Computing liberates the cloud user from providing their own physical resources required for carrying out data processing. Instead, the user is able to utilise scalable cloud resources on demand when making use of a cloud provider's infrastructure or even software. This may lead to financial advantages for the cloud user. It is questionable, though, if financial advantages are sufficient to outweigh the data subject's interest concerning comprehensive data protection. [233] This interest is based on a European fundamental right, Article 7 and 8 CFR. [234]

In a recent decision, the ECJ was required to evaluate a similar conflict of interests: a Spanish citizen demanded that Google was obliged to remove personal data within search results and cease from making the relevant search results available to the public. [235] The Spanish citizen argued that Google could no longer base the processing of the data on the legal grounds of Article 7 lit (f). The ECJ emphasized the strong position of the data subject stating that, in this case, merely the economic interest of Google would not be able to justify the processing. [236] The court held that as a rule the rights of the data subject override the economic interests of the operator of the search engine. [237] Although the decision was not concerned with a balance of interests between a cloud user and an affected data subject, but rather with a data subject who was facing a disadvantageous Google search result (linking information to him), the ECJs reasoning can also be applied to Cloud Computing, as well. Therefore, processing on the grounds of Article 7 lit (f) should not only be justifiable through reasons of financial advantage for the cloud user.

[233] Against a saving of costs as a justification for processing *Nägele/Jacobs*, ZUM 2010, 281 (290); *Niemann/Paul*, K&R 2009, 444 (449) on the other hand recognizing a saving of costs as a justification; principally recognizing financial aspects as an possible justification, but only if it would be unreasonable for the controller to waive the processing *Hoeren*, in: Roßnagel, Handbuch Datenschutzrecht, Chapter 4.6, rec. 31, see also *Spindler/Nink*, in: Spindler/Schuster, Recht der elektronischen Medien, par. 28 BDSG, rec. 6

[234] Charter of Fundamental Rights of the European Union, Official Journal of the European Communities, 2000/C 364/01 of the 18.12.2000, available at: `http://www.jura.uni-wuerzburg.de/fileadmin/02120300/_temp_/Abbreviations.pdf`.

[235] *ECJ*, Judgment from the 13th May 2014 in Casec-131/12 – Google Spain SL/Google Inc. v AEPD/Mario Costeja Gonzalez.

[236] *ECJ*, Judgment from the 13th May 2014 in Casec-131/12 – Google Spain SL/Google Inc. v AEPD/Mario Costeja Gonzalez. par. 81

[237] *ECJ*, Judgment from the 13th May 2014 in Casec-131/12 – Google Spain SL/Google Inc. v AEPD/Mario Costeja Gonzalez. par. 97

2.4.2.2 Legal Permissions in the GDPR

The proposal for a GDPR includes permissions for the processing of personal data that function as exceptions from the general prohibition referred to in 2.1.2. Besides the previous given consent by the affected person (see 2.4.2.3), Article 6 GDPR mentions five other exceptions (wording of the LIBE-Proposal):

> "[...]
>
> (b) processing is necessary for the performance of a contract to which the data subject is party, or in order to take steps at the request of the data subject prior to entering into a contract;
>
> (c) processing is necessary for compliance with a legal obligation to which the controller is subject;
>
> (d) processing is necessary in order to protect the vital interests of the data subject;
>
> (e) 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;
>
> (f) processing is necessary for the purposes of the **legitimate interests** pursued by the controller, or in case of disclosure by the 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. (emphasis added)

Essentially, the GDPR allows for data processing only if the affected party consents to it or if the data processing is necessary for legitimate purposes pertaining to the processor. For Cloud Computing, the same problems as described in 2.4.2.1 exist for a legally admitted processing without consent. It has been criticized that lit. (b) only covers contractual claims and does not include statutory claims. [238] Nevertheless, lit. (b) includes the "performance of a contract", without a restriction to claims. Moreover, lit. (f) covers all legitimate interests, in case they are not overridden by the data subjects interests; thus, data processing in order to enforce a statutory claim could be lawful without consent of the affected person under certain circumstances. Although, if the processing for direct marketing purposes is permitted according to lit. (f), the data subject is able to object to the processing at any time, free of charge, and without any further justification, Article 21 Par. 2 GDPR. According to Article 21 Par. 1 GDPR the data subject has the right to object to the processing of their data with the result that "the controller shall no longer process the personal data unless the controller demonstrates compelling legitimate grounds for the processing which override the interests, rights and freedoms of the data subject or for the establishment, exercise or defense of legal claims". This extensive right to object does not exist if the processing is based on lit (b), which might in turn cause lit (b) to be the more reliable rationale for processing.

2.4.2.3 Informed Consent and Cloud Computing

2.4.2.3.1 Data Protection Directive

In case the DPD is applicable, the safest way to ensure compliance with the Data Protection Directive's national acts of implementation is to ask the data subject for his or her explicit consent. Cited by the Directive in Article 7, consent is the first out of seven legal grounds for personal data processing.

[238] *Berg*, PinG 2013, 69 (70).

According to Article 8 of the Directive, consent needs to be given explicitly for processing special categories of data. Some Member States view consent as a preferred ground for lawfulness, whereas others view it as one of six options. Every other legal basis for data processing requires a necessity-test. In contrast, the data subject's consent allows the data processor to go beyond what is necessary for their purposes; [239] in other words, the data processor is not bound by a strict proportionality test under these circumstances. [240] However, as noted already, the DPD treats specific sensitive data in an intensified manner, such as data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and data concerning health or sex life.

Consent in the sense of the law is only effective when it is informed and given freely, and unambiguously. Informed consent implies that the data subject has been given certain information before data is processed, including the recipients or categories of recipients of the data (Article 10 (c) Data Protection Directive). [241] It also should additionally be made clear to the data subject whether data will be transferred to a non-EU-state. [242]

These requirements of ex-ante information and transparency lead to difficulties. For example: in case data is to be computed in a cloud, it might be hard to tell when the data will be transferred to a server (to which server?) and in which state this server will be operated. [243] Due to the scalability of cloud computing, the method of storage and the "division of labour" amongst the different servers might be 'decided' by automated programs and could change within seconds. [244] A distinction has to be made between two scenarios. In the first one, the cloud user is the data subject himself (e.g. the user of an online e-mail service, like Gmail). In the other scenario, the cloud user is outsourcing data of a third party (e.g. a company handles its clients' data with a cloud solution). In the first scenario, only one data subject has to give consent to the data processing, which can be done before the user subscribes to the cloud service. In the other scenario, the user would need the consent of every single one of his clients (cascade of consent). This can be done when a new client and the cloud user make their first contractual agreement regarding whatever service the user is offering, but it becomes nearly impossible for old clients, as every one of those must be contacted, given time to react, and give his consent. Implementing the declaration of consent in the general contract terms and conditions might be valid if the client has to accept them actively, but changing existing terms and conditions and informing old clients does not provide a given consent. If consent is given by accepting terms and conditions, legal requirements regarding consumer protection law have to be met, as well. [245]

To ensure compliance with the data protection law, consent would not be the best solution in such a case. [246] For a cloud provider, it would be useful to make the storage of data scalable by location, so that the user can choose certain servers to be used for their computation. [247] This way, the data subject can be informed specifically about the location of their data. The same principle applies to the provider's sub-contractors: the user should choose which subcontractor they will use for the specific computation, thus informed consent is safeguarded. Nevertheless, this might be impossible for cloud computing services using resources of other cloud providers. For instance, SaaS provider, Dropbox,

[239] *Art. 29-Working Party*, Opinion 15/2011, WP 187, 7f.

[240] *Nägele/Jacobs*, ZUM 2010, 281 (290); *Rath/Rothe*, K&R 2013, 623 (624).

[241] *Taeger*, in: Taeger/Gabel, BDSG, par. 4a, recital 30; *Nord/Manzel*, NJW 2010, 3756 (3757).

[242] *Simitis*, in: Simitis, Bundesdatenschutzgesetz, par. 4a, recitals 70 ff.

[243] *Nägele/Jacobs*, ZUM 2010, 281; *Schultze-Melling*, in: Taeger/Gabel, BDSG, par. 9, recital 104.

[244] *Millard*, Cloud Computing, Chapter 1.1, 1.2; *Funke/Wittmann*, ZD 2013, 221 (222).

[245] *Spindler*, GRUR-Beilage 2014, 101.

[246] *BITKOM*, Leitfaden Cloud Computing, p. 51; *German Federal Office for Information Security Technology*, Safety Recommendation for Cloud Computing Providers, p. 73; *Brennscheidt*, Cloud Computing, p. 152; *Art. 29-Working Party*, Opinion 15/2011, WP 187, 12.

[247] For an example such a service is provided by Amazon Web Services, available at: `http://aws.amazon.com/de/ec2/pricing/effective-april-2014/`

builds their service on IaaS by Amazon's S3 (a double 'layer'). Even an SaaS built on a PaaS lying on an IaaS, itself, is possible (e.g. Facebook apps on Heroku on Amazon). [248] In such cases, cloud users will not necessarily know in which data centers, or even countries, their data is stored or with whom their provider has a subcontractor relationship. In fact, the providers may not even be aware themselves. [249]

2.4.2.3.2 Informed Consent and Obligation of Transparency under the GDPR

Article 13 GDPR extends the approach of the DPD concerning transparency for data subjects (and also goes beyond existing national laws, like in Germany) [250] by specifying to the data subject the information that has to be supplied prior to the collection of data. For the purposes of cloud computing, these obligations to inform raise a lot of issues. In order to get an idea of the upcoming problems, we have to take a closer look at the required information of Articles 13 and 14 GDPR. Article 13 applies if personal data relating to a data subject are collected from the data subject whereas Article 14 applies if personal data have not been obtained from the data subject. According to Article 13 Par. 1

> "(. . .) the controller shall provide the data subject with at least the following information:
>
> (a) the identity and the contact details of the controller and, where applicable, of the controller's representative;
>
> (b) the contact details of the data protection officer, where applicable;
>
> (c) the purposes of the processing for which the personal data are intended as well as the legal basis for the processing;
>
> (d) where the processing is based on point (f) of Article 6(1), the legitimate interests pursued by the controller or by a third party;
>
> (e) the recipients or categories of recipients of the personal data, if any;
>
> (f) where applicable, the fact that the controller intends to transfer personal data to a third country or international organization and the existence or absence of an adequacy decision by the Commission, or in the case of transfers referred to in Article 46 or 47, or the second subparagraph of Article 49(1), reference to the appropriate or suitable safeguards and the means to obtain a copy of them or where they have been made available;

To complete the provisions concerning transparency for data subjects in Par. 1, further provisions have been added to Article 13 Par. 2 - 4 GDPR:
In addition to the information referred to in Par. 1, the controller shall, at the time when personal data are obtained, provide the data subject with the following further information necessary to ensure fair and transparent processing:

> (a) the period for which the personal data will be stored, or if that is not possible, the criteria used to determine that period;
>
> (b) the existence of the right to request from the controller access to and rectification or erasure of personal data or restriction of processing concerning the data subject or to object to processing as well as the right to data portability;

[248] *Millard*, Cloud Computing, Chapter 3.2.

[249] *ComputerWorldUK Cloud Vision blog*, Cloud computing and EU data protection law, Part one: Understanding the international issues, available at: `http://blogs.computerworlduk.com/cloud-vision/2011/09/cloud-computing-and-eu-data-protection-law/index.htm`.

[250] *Jaspers*, DuD 2012, 571 (572).

> (c) where the processing is based on point (a) of Article 6(1) or point (a) of Article 9(2), the existence of the right to withdraw consent at any time, without affecting the lawfulness of processing based on consent before its withdrawal;
>
> (d) the right to lodge a complaint with a supervisory authority;
>
> (e) whether the provision of personal data is a statutory or contractual requirement, or a requirement necessary to enter into a contract, as well as whether the data subject is obliged to provide the personal data and of the possible consequences of failure to provide such data;
>
> (f) the existence of automated decision-making, including profiling, referred to in Article 22(1) and (4) and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject.

Moreover, where the controller intends to further process the personal data for a purpose other than that for which the personal data were collected (see above 2.4.2.2), according to Par. 3 he or she shall provide the data subject prior to that further processing with information on that other purpose and with any relevant further information as referred to in Par. 2. However, these obligations stipulated in Par 1-3 “shall not apply where and insofar as the data subject already has the information”.

It is evident that due to the variety of data processing procedures and sub-providers in the cloud, it is nearly impossible to provide sufficient information to the data subject. For instance, the recipients of the personal data cannot be realistically be identified in a cloud in advance, as the storing and processing depends upon the (global and dispersed) available capacities. The same is true for the transfer of data to a third country - which cannot be easily assessed in advance (see 2.4.2.3). If a controller intends to collect data by using a cloud service in order to process the data, it will be even more important for them that the data is not considered “personal data” under the GDPR.

According to Article 13 GDPR, the controller has to provide the information at the time when personal data are obtained which also includes the collecting of data based on an explicit legal permission, as provided by Article 6. In contrast, the DPD simply requires the controller to provide such information to gain informed consent. The GDPR, on the other hand, requires that such information is also provided when data is collected, on the grounds of legal permission.

The cloud user (controller) can comply with these obligations by choosing a cloud provider who enables them to determine which servers in what country will be used to offer the cloud service. [251]

Whereas the DPD only requires a freely given consent - particularly informed consent - the GDPR demands far more requirements from a controller.

Article 4 No. 11 GDPR intensifies the requirements for a valid consent by demanding an “specific, informed and unambiguous” consent [252] and moreover “a statement or (. . .) a clear affirmative action.” Thus, the possibility of hiding statements in terms and conditions or using implied consent would no longer be possible. [253] The users will normally have to ’opt-in’. This is highlighted by Recital 32 S. 3 GDPR which states that “silence, pre-ticked boxes or inactivity should 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.” Recital 32 S. 2 GDPR provides further examples for a valid consent, such as “ticking a

[251] The determination of the servers used is possible e.g. if Amazon is chosen as a cloud provider, see `http://aws.amazon.com/about-aws/global-infrastructure/regional-product-services/`.

[252] An ”explicit” consent - as stipulated in the Commission’s and in the Parliament’s proposals - has not been included into the GDPR’s final version. However, if special categories of personal data - such as sensitive data - are processed an explicit consent according to Article 9 Par. 2 lit. (a) GDPR is still required.

[253] *Hartung/Storm*, in: Hilber, Handbuch Cloud Computing, Teil 4, recital 66.

box when visiting an internet website", "choosing technical settings for information society services"- which some academics interpret as a way of giving valid consent by using the technical settings of a browser [254] - and 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". The GDPR's minimum requirements for informed consent are that "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" (see Recital 42 S. 3 GDPR). Moreover, according to Recital 32 S. 4, 5 GDPR the data subject's consent "should cover all processing activities carried out for the same purpose or purposes. When the processing has multiple purposes, consent should be given for all of them."

In accordance with Article 6 Par. 1 (a) GDPR, the processing of personal data shall be legitimated through unambiguous consent only if this consent refers to specific and defined purposes. Nevertheless, there are some exceptions to the principle of strict purposes (see above 2.4.2.2).

A data subject's consent will not be regarded as freely given if he or she "has no genuine or free choice or is unable to refuse or withdraw consent without detriment" (Recital 42 S. 4 GDPR). Regarding the protection of the free choice of a data subject Article 7 Par. 4 GDPR establishes a "prohibition of coupling":

> "When assessing whether consent is freely given, utmost account shall be taken of whether, inter alia, the performance of a contract, including the provision of a service, is conditional on consent to the processing of personal data that is not necessary for the performance of that contract." (emphasis added)

Moreover, according to Recital 43 S. 2 GDPR "(c)onsent is presumed not to be freely given if it does not allow separate consent to be given to different personal data processing operations despite it being appropriate in the individual case, or if the performance of a contract (. . .) is dependent on the consent despite such consent not being necessary for such performance." According to this, consent will in many of the cases (cf. the wording of Article 7 GDPR "utmost account") be invalid if it is linked with a consent to a contract which is not necessary for the original purpose of the consent. Additionally, 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" (see Recital 43 S. 1 GDPR) this imbalance leads to the assumption that the consent of the data subject was not freely given and thus has to be considered as invalid.

Article 7 Par. 3 GDPR provides the data subject with the right to withdraw his or her consent at any time without any further requirements and this right has to be articulated prior to the data subject giving consent. Furthermore, Article 8 Par. 1 Sentence 2 GDPR states that the processing of personal data of a child below the age of 16 years shall only be lawful if and to the extent that consent is given by the holder of parental responsibility. Par. 2 states that "[t]he controller shall make reasonable efforts to verify in such cases that consent is given or authorised by the holder of parental responsibility over the child, taking into consideration available technology."

However, Article 7 of the GDPR does not demand a written form for a consent, thus signifying that electronic consent and even an oral statement is sufficient.

The obligations to inform are flanked by the provision in Article 12 Par. 7 GDPR that requires the data controller to provide standardized and easily legible information in order to give a meaningful

[254] *Härting, Datenschutz-Grundverordnung*, 2016, Recital 364; different opinion Spindler, DB 2016, 937 (940) who doubts whether default settings of a browser are sufficient for a valid consent because otherwise the regulations in Article 25 Par. 5 of the proposal of the Council for a GDPR which suggested such a regulation had been included into the final version of the GDPR which however is not the case.

overview of the intended processing. In this context the adaption of a "One-Pager" [255] pattern is discussed which could enable the data controller to present its essential processing information on a well-structured single page. Although the "One-Pager" seems to be both standardized and legible, it leads to an (optical) reduction of the given information to the extent that it might not cover all of the intended processing. Furthermore, the implementation of visualised concepts as the labelling in the food industry [256] for instance, could increase the legal certainty of data processing by providing more transparency to the data subject. [257] Concerning the information of Article 13 GDPR, the information has to be specified according to the individual circumstances of the data subject; for instance, information about the national competent supervising authority or about options to file a complaint.

Finally, the lawfulness of consents under the GDPR which were obtained under the conditions of the legal framework of the DPD is unclear. Regarding this issue the German data protection authorities ("DPAs"), the so called "circle of Düsseldorf (Düsseldorfer Kreis)", published a non-binding opinion. Recital 171 S. 2 GDPR stipulates that "(w)here processing is based on consent pursuant to Directive 95/46/EC, it is not necessary for the data subject to give his or her consent again if the manner in which the consent has been given is in line with the conditions of this Regulation". According to the DPAs, validly given consents under the DPD generally fulfill these requirements. The DPAs state that regarding this the information requirements of Article 13 GDPR do not have to be fulfilled, since these requirements do not belong to the conditions referred to in Recital 171. However, the DPAs point out that the conditions of Article 7 Par. 4 and Article 8 Par. 1 GDPR (see above) have to be fulfilled. If these two conditions are not met in the existing consents, the consents will not continue to be lawful under the GDPR. [258]

2.4.3 Data transfer to third Countries

2.4.3.1 The DPD

Unless the data subject consents or the provisions of the DPD expressly permit it, transferring data to a 'third country' (a state not within the EU or the EEA) is principally forbidden. [259] The same problems mentioned above can also occur in connection with the data subject's consent to data processing in a cloud when the data subject has to agree to a transfer in an unsafe country. [260] Hence, there is a difference between the legal permission to process data and the legal permission to transfer the data into a third country. The data transfer is only rendered if both requirements are fulfilled separately.

One of the main exceptions refers to the 'adequate level of protection' in the third country, Article 25 of the Directive. [261] An adequate level of protection assumes that the data protection standards in the respective country are comparable to European standards. This has to be officially acknowledged by the European Commission as has been the case for the following countries: Andorra, Argentina, Australia, Canada, Switzerland, Faeroe Islands, Guernsey, the State of Israel, the Isle of Man, Jersey,

[255] Pressemitteilung des Bundesministeriums der Justiz und fr Verbraucherschutz, available at: http://www.bmjv.de/SharedDocs/Pressemitteilungen/DE/2015/11192915_Vorstellung_OnePager.html.

[256] Seehttp://www.vzbv.de/sites/default/files/mediapics/was_ist_die_ampel.pdf.

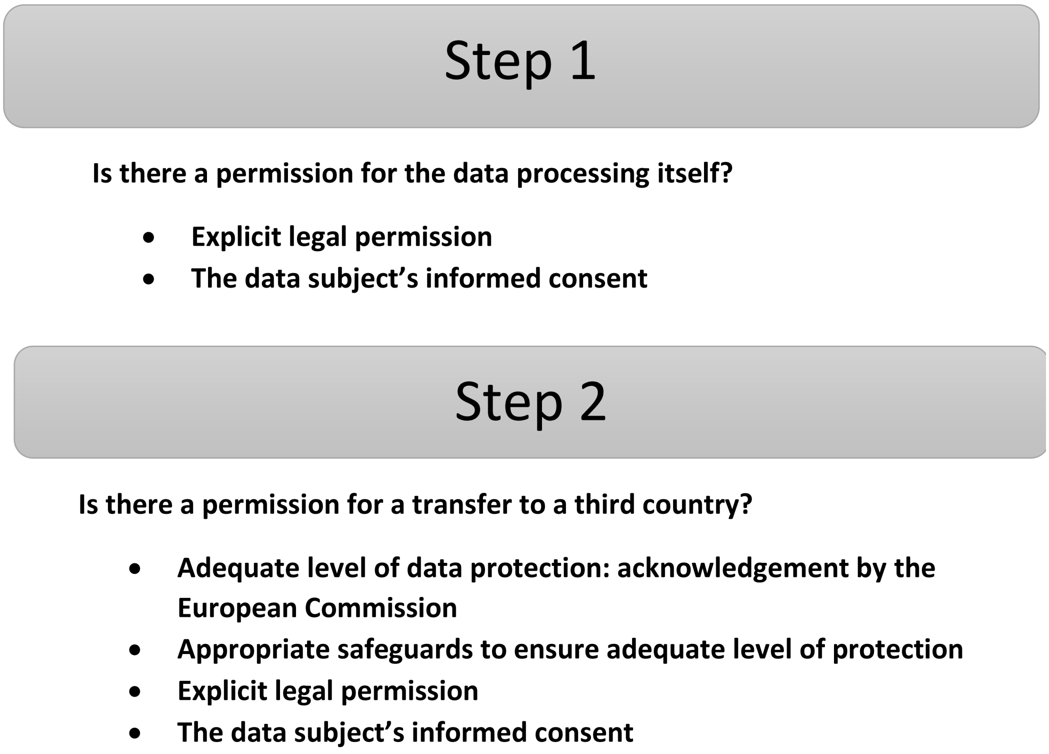
[257] *Pollmann/Kipker*, DuD 2016, 378 (379).

[258] Düsseldorfer Kreis, decision of 13th/14th September 2016, available at: https://www.lda.bayern.de/media/dk_einwilligung.pdf.

[259] In detail *Art. 29-Working Party*, Working Document on Transfers of personal data to third countries: Applying Article 26 (2) of the EU Data Protection Directive to Binding Corporate Rules for International Data Transfers, WP 74.

[260] See also *Brennscheidt*, Cloud Computing, p. 175.

[261] *Hon/Millard*, Data Export in Cloud Computing - How can Personal Data be Transferred outside the EEA?, p. 5.



Step 1

Is there a permission for the data processing itself?

- Explicit legal permission
- The data subject's informed consent

Step 2

Is there a permission for a transfer to a third country?

- Adequate level of data protection: acknowledgement by the European Commission
- Appropriate safeguards to ensure adequate level of protection
- Explicit legal permission
- The data subject's informed consent

Figure 2.5: Data transfers to third countries

New Zealand, and Eastern Republic of Uruguay. [262] Being of particular relevance for Cloud Computing, the USA has not been acknowledged. [263] However, the Safe Harbor agreement between the EU and the USA provides a possibility for American companies to comply with the DPD when transferring personal data to the USA. [264] The companies can certify themselves and opt into the program by signing a declaration of accession and by publishing a privacy statement of the US Department of Commerce. [265] The USA does not provide an adequate level of data protection in accordance with Article 25 Par. 6 DPD, therefore Safe Harbor has been negotiated outside the scope of Article 26 DPD as an international treaty between the EU and the USA, the principles of the agreement are based on Article 25 Par. 1 and 2 DPD. [266]

Nevertheless, this changed crucially since the European Court of Justice decided as a result of the case Maximillian Schrems versus Data Protection Commissioner that the decision of the Commission regarding the adequate level of data protection in the US is invalid. [267]

According to the ECJ a system of self-certification does not contradict the requirement laid down in Article 25 Par. 6 DPD. However, the reliability of such a system is at stake, in particular that the

[262] All decisions by the European Commission regarding the acknowledgment of third countries are available at: `http://ec.europa.eu/justice/data-protection/document/international-transfers/adequacy/index_en.htm`.

[263] No such decision has been made by the commission; *Gabel*, in: Taeger/Gabel, BDSG, par. 4b, recital 23; *BITKOM*, Leitfaden Cloud Computing, p. 53.

[264] The Safe Harbor Principles are available at: http://export.gov/safeharbor/; Hon/Millard, Data Export in Cloud Computing - How can Personal Data be Transferred outside the EEA?, p. 15; a summary of the essential rules of the Safe Harbor Principles is provided by *Wisskirchen*, CR 2004, 862 (864 f.).

[265] *Hartung/Storm*, in: Hilber, Handbuch Cloud Computing, Teil 4, recital 237.

[266] *Savin*, EU Internet Law, p. 204; v. d. *Bussche*, in: Plath, BDSG, par. 4b, recital 30. At the moment, there are still ongoing negotiations between the EC and the USA about a new agreement.

[267] *ECJ*, decision of 06/10/2015, Case C362/14 - Request for a preliminary ruling of the ECJ from the High Court (Ireland), Maximillian Schrems v Data Protection Commissioner and Digital Rights Ireland Ltd.

concept of adequacy is (also) essentially founded on the establishment of effective detection and supervision mechanisms concerning any infringements upon the rules, thus ensuring the protection of fundamental rights (recital 81). The Safe Harbor principles are applicable solely to self-certified US organizations whereas US public authorities are not required to comply with them (recital 28). The court states that "national security, public interest, or law enforcement requirements have primacy over the Safe Harbor principles, primacy pursuant to which self-certified United States organisations receiving personal data from the European Union are bound to disregard those principles without limitation where they conflict with those requirements and therefore prove incompatible with them" (recital 86). Thus, the Safe Harbor Agreement enables the US authorities to infringe in European fundamental rights (c.f. recital 87). Furthermore, in recital 89 the court criticizes the lack of effective legal protection against interferences of that kind and refers to the opinion of European Court of Justice's advocate general Yves Bot [268] in its recital 204 according to which "the private dispute resolution mechanisms and the FTC, owing to its role limited to commercial disputes, are not means of challenging access by the US intelligence services to personal data transferred from the European Union." The ECJ states that "legislation not providing for any possibility for an individual to pursue legal remedies in order to have access to personal data relating to him, or to obtain the rectification or erasure of such data, does not respect the essence of the fundamental right to effective judicial protection, as enshrined in Article 47 of the Charter which requires everyone whose rights and freedoms guaranteed by the law of the European Union are violated to have the right to an effective remedy before a tribunal in compliance with the conditions laid down in that article" (recital 95). Nevertheless, the Commission did not state in the Safe Harbor Agreement that the USA ensures an adequate level of protection by reason of its domestic law or its international commitments (recital 97). Thus, "it is to be concluded that Article 1 of Decision 2000/520 fails to comply with the requirements laid down in Article 25(6) of Directive 95/46, read in the light of the Charter, and that it is accordingly invalid" (recital 98). Furthermore, the agreement shall be invalid, "because the national supervisory authorities must be able to examine, with complete independence, any claim concerning the protection of a person's rights and freedoms in regard to the processing of personal data relating to him. That is in particular the case where, in bringing such a claim, that person raises questions regarding the compatibility of a Commission decision adopted pursuant to Article 25(6) (...) (recital 99). However, Article 3 Par. 1 of the Safe Harbor decision lays down specific rules regarding the powers available to the national supervisory authorities in light of a Commission finding relating to an adequate level of protection (recital 100) and is consequently invalid, as well as the decision in its entirety (recital 105). As a consequence of the Safe Harbor decision, the Commission and the US government established a new legal framework to regain legal certainty when transferring personal data to the United States. After having consulted several opinions [269] the Commission adopted a new agreement, the EU-US Privacy Shield Framework, which came into effect on 12 July 2016. [270] Similar to Safe Harbor, the EU-US Privacy Shield uses a system of self-certification for companies. [271] Beside strong obligations

[268] Opinion of Advocate General *Yves Bot*, delivered on 23/09/2015, Case C362/14 - Request for a preliminary ruling of the ECJ from the High Court (Ireland), Maximillian Schrems v Data Protection Commissioner.

[269] See e.g. Art. 29-Working Party, Opinion 01/2016 on the EU - U.S. Privacy Shield draft adequacy decision, WP 238, available at: `http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2016/wp238_en.pdf`.

[270] Commission Implementing Decision (EU) 2016/1250 of 12 July 2016 pursuant to Directive 95/46/EC of the European Parliament and of the Council on the adequacy of the protection provided by the EU-U.S. Privacy Shield, Official Journal of the European Union L 207/1-112, available at: `http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016D1250&from=EN`; see also European Commission, Guide to the EU-U.S. Privacy Shield, 2016, available at: `http://ec.europa.eu/justice/data-protection/document/citizens-guide_en.pdf`.

[271] A list of all companies certified under the EU-US Privacy Shield can be found at `https://www.privacyshield.gov/list`.

for companies handling data and clear safeguards and transparency obligations on US government access, the EU-US Privacy Shield requires the effective protection of individual rights. The warranty and functioning of these principles will be monitored by an annual joint review mechanism conducted by the European Commission and the US Department of Commerce. [272] However, as well as the former Safe Harbor agreement, the new transatlantic framework provides legal uncertainty in terms of contents and proceedings which relies to a certain extent on the chosen concept (the EU-US Privacy Shield consists of diverse correspondence letters). [273] Additionally, the EU-US Privacy Shield has been established with regards to the provisions of the DPD - however, the compliance of the agreement with the upcoming GDPR is uncertain. [274] Regarding cloud-computing based data transfers, it seems to be unlikely that the EU-US Privacy Shield increases the level of data protection for data subjects. Especially in the public area, the new agreement does not provide detailed provisions to strengthen the enforcement of rights of the affected data subject against the public authorities. The new agreement rather guarantees the US authorities the access to personal data as the collection of data is only restricted to a required extent. Thus, the legal situation has not significantly improved since Safe Harbor. [275] However, one potential solution would involve acquiring the consent of the data subjects for the data transfer, which in most of the cases would not be possible.

Besides the officially acknowledged countries mentioned above (where no explicit consent from the user is needed), the data controller who wishes to transfer data in other countries may use other forms of justification provided by the DPD. In general, this is possible, if the controller adduces evidence of adequate safeguards with respect to the protection of the data subject's rights, Article 26 Par. 2 of the Directive. Those safeguards can be based upon appropriate standard contractual clauses which the EU Commissions has acknowledged regarding processors in third countries [276] between the controller and the entity receiving the data, ensuring an adequate level of protection. Those clauses are used to establish rules for the third country party that displays the same standard of protection as the EU does for data subject's rights. Yet, the benefit of a lawful transfer to the third country only exists if the clauses acknowledged by the Commission are used exactly how the Commission provided them and without alteration. [277]

However, it is moreover doubtful if standard contractual clauses are considered to be lawful. In the light of the reasoning of the ECJ's decision on the Safe Harbor agreement, it seems likely that also standard contractual clauses that have been set by the EU Commission will be unlawful. [278] As a consequence thereof, the EU Commission recently presented two draft decisions amending the existing adequacy decisions and the decisions on standard contractual clauses in order to cure the

[272] European Commission - Press release, available at: `http://europa.eu/rapid/press-release_IP-16-2461_en.htm`.

[273] *Weichert*, ZD 2016, 209 (214).

[274] *Boerding*, CR 2016, 431 (440).

[275] *Boerding*, CR 2016, 431 (438).

[276] See also *Art. 29-Working Party.*, Opinion 03/2009, WP 161; Standard Contractual Clauses I, Commission Decision of 15th June 2001 on standard contractual clauses for the transfer of personal data to third countries, under Directive 95/46/EC, C(2001) 1539 (2001/497/EC), available at: `https://www.datatilsynet.no/Global/04_skjema_maler/EUs%20standardkontrakter1_ENG.pdf`; Standard Contractual Clauses II, Commission Decision of 27th December 2004 amending Decision 2001/497/EC as regards the introduction of an alternative set of standard contractual clauses for the transfer of personal data to third countries, C(2004) 5271 (2004/915/EC), available at: `http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2004:385:0074:0084:en:PDF`; Commission Decision 2010/87/EU of 05.02.2010 on Standard Contractual Clauses for Data Processors established in Third Countries, available at: `http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:039:0005:0018:EN:PDF`.

[277] *Gola/Klug/Körffer*, in: Gola/Schomerus, Bundesdatenschutzgesetz, par. 4c, recital 14; *Spindler/Nink*, in: Spindler/Schuster, Recht der elektronischen Medien, par. 4c BDSG, recital 20.

[278] See the commended proceedings at the Irish High Court involving Facebook and Maximilian Schrems, available at: `http://www.europe-v-facebook.org/PA_MCs.pdf`.

illegality that follows from the findings in the ECJ's Schrems ruling. [279] Nonetheless, as long as there is no judgment of the ECJ on this regard, for a cloud user who wishes to transfer data to a cloud provider within a third country, the standard contractual clauses would only be useful if the cloud provider agrees to those exact clauses. It seems unlikely that a cloud provider in contractual agreement with many cloud users would alter these existing contracts, but would rather agree to the standard contractual clauses. Another way of providing adequate safeguards are the so-called Binding Corporate Rules (BCR). Other than the standard contractual clauses, BCR are not mentioned explicitly in the Directive. Nevertheless, Article 26 Par. 2 is not exhaustive, which suggests that appropriate safeguards might be other measures than the explicitly mentioned standard contractual clauses; they are only an example among other possible safeguards. [280] BCR are supposed to ensure that there is an adequate level of data protection for data transfers within a corporation, regardless of the countries the corporation might be seated in

> "The rules must apply generally throughout the corporate group, irrespective of the place of establishment of the members, or the nationality of the data subjects whose personal data is being processed, or any other criteria or consideration." [281]

The BCR have to be binding or legally enforceable and should be regarded as "sufficient safeguards" within the context of Article 26 Par. 2 DPD. They are meant to be used by multinational companies to allow international data transfers. [282] There are no model BCR provided by the Art. 29-Working Group or the Commission, such as in the case of standard contractual clauses. However, the Art. 29-Working Group proposed crucial elements of BCR and how these rules might be structured in a single document. [283] BCR have to be acknowledged by a supervisory authority in an EU Member State. In case of such an acknowledgement, authorities of most EU-Member States acknowledge BCR automatically, thus creating some form of European passport (notwithstanding the fact that the DPD does not contain such a procedure). [284] In some specific cases, BCR may be used for cloud computing-related data transfers, however, they will be restricted to internal data transfers across borders. [285] On the other hand, most cloud related data transfers to third countries will not be within a corporation but rather effectuated in a cloud, thus transferring data from a cloud user to a cloud provider. Therefore, BCR do not provide a general solution for cloud computing related to third-country transfers. [286] Nevertheless, following the decision of the ECJ regarding Safe Harbor it is doubtful whether standard contractual clauses or BCR can be seen as a lawful way to transfer data to the USA, because they also don't prevent US authorities to access personal data transferred to the USA.

[279] See the Summary record of the 72nd meeting of the Committee on the Protection of Individuals with regard to the Processing of Personal Data (Article 31 Committee), 3 October 2016, available at: `http://ec.europa.eu/transparency/regcomitology/index.cfm?do=search.documentdetail&cwD7uvTdufhI1m2L+QpBt4y9MhSLwjviGkXX0DmrSK4XV3U4/r7rgJvJWdYwELHg`.

[280] *Art. 29-Working Party*, Working Document on Transfers of personal data to third countries: Applying Article 26 (2) of the EU Data Protection Directive to Binding Corporate Rules for International Data Transfers, WP 74, p. 6.

[281] *Art. 29-Working Party*, Working Document on Transfers of personal data to third countries: Applying Article 26 (2) of the EU Data Protection Directive to Binding Corporate Rules for International Data Transfers, WP 74, p. 8.

[282] *Art. 29-Working Party*, Working Document on Transfers of personal data to third countries: Applying Article 26 (2) of the EU Data Protection Directive to Binding Corporate Rules for International Data Transfers, WP 74, p. 8.

[283] C.f. *Art. 29-Working Party*, Working Document Setting up a framework for the structure of Binding Corporate Rules, WP 154.

[284] *Brennscheidt*, Cloud Computing, p. 173.

[285] *Niemann/Paul*, K&R 2009, 444 (449).

[286] *Brennscheidt*, Cloud Computing, p. 174.

2.4.3.2 The GDPR

Concerning the transfer of data to companies/data processors located outside the EU, the Regulation adheres to the former approach of the DPD. [287] The GDPR also demands the two steps necessary for a lawful transfer, as mentioned above (2.3.2.6). The first step refers to the permission to process the personal data. The second one concerns the transfer of data to a third country, thus safeguarding an adequate level of protection (comparable to the European level), which is crucial for any transmission. However, as a result of the ECJ decision on Safe Harbor (see above 2.4.3.1) new principles have been implemented in the GDPR to determine whether or not the receiving third country is granting a sufficient level of data protection, Article 45 Par. 2 GDPR. The influence of the judgment on the final version of the GDPR becomes obvious by pointing out some of the amendments, such as the access of public authorities to personal data, as well as the implementation of such legislation or effective and enforceable data subject rights and effective administrative and judicial redress for the data subjects whose personal data are being transferred. [288] Recital 104 GDPR underlines that there must be an adequate level of protection which is "essentially equivalent" to that ensured within the EU. Moreover, once the equivalent level is approved, the Commission shall monitor and reconsider the functioning of the decision on the level of protection in the third country. Therefore, the implementing act shall provide for a "periodic review, at least every four years, which shall take into account all relevant developments (. . .)" (Article 45 Par. 3 GDPR). The instruments which a data processor can use to comply with these requirements remain basically the same. The transmission of personal data to third countries can be based on:

- an acknowledgement of adequacy by the EU Commission (Article 41 GDPR), or
- 'binding corporate rules' (Article 42 Par. 2 a and Article 43 GDPR),
- standard data protection clauses adopted by the Commission (Article 46 Par. 2 (c))
- standard data protection clauses adopted by a supervisory authority and approved by the Commission (Article 46 Par. 2 (d) GDPR), or
- an approved code of conduct pursuant to Article 40 (Article 46 Par. 2 (e)) or an approved certification mechanism pursuant to Article 42 (Article 46 Par. 2 (f)) both together with binding and enforceable commitments of the controller or processor in the third country to apply the appropriate safeguards, including with regards to data subjects' rights (see 2.3.2.4);

Concerning the benchmarks and relevant criteria which the EU Commission will use for acknowledgement of an adequate level, the wording of Article 45 Par. 2 GDPR requires the Commission to give consideration to the following elements:

> " (a) the rule of law, respect for human rights and fundamental freedoms, relevant legislation, both general and sectoral, including concerning public security, defence, national security and criminal law and the access of public authorities to personal data, as well as the implementation of such legislation, data protection rules, professional rules and security measures, including rules for the onward transfer of personal data to another third country or international organisation which are complied with in that country or international organisation, case-law, as well as effective and enforceable data subject rights and

[287] *Nebel/Richter*, ZD 2012, 407 (412).

[288] Cf. *Bergt*, Der Einfluss des Safe-Harbour-Urteils auf den Entwurf der Datenschutz-Grundverordnung, available at: `http://www.cr-online.de/blog/2016/01/04/der-einfluss-des-safe-harbor-urteils-auf-den-entwurf-der-datenschutz-grundverordnung/`.

> effective administrative and judicial redress for the data subjects whose personal data are being transferred;
>
> (b) the existence and effective functioning of one or more independent supervisory authorities in the third country or to which an international organization is subject, with responsibility for ensuring and enforcing compliance with the data protection rules, including adequate enforcement powers, for assisting and advising the data subjects in exercising their rights and for cooperation with the supervisory authorities of the Member States; and
>
> (c) the international commitments the third country or international organization concerned has entered into, or other obligations arising from legally binding conventions or instruments as well as from its participation in multilateral or regional systems, in particular in relation to the protection of personal data."

In addition recital 104 of the GDPR states that:

> "In line with the fundamental values on which the Union is founded, particularly the protection of human rights, the Commission should, in its assessment of the third country, or of a territory or specified sector within a third country, take into account how a particular third country respects the rule of law(...)".

Moreover, the GDPR now includes that:

> "The third country should offer guarantees ensuring an adequate level of protection essentially equivalent to that ensured within the Union, in particular where personal data are processed in one or several specific sectors. In particular, the third country should ensure effective independent data protection supervision and should provide for cooperation mechanisms with the Member States' data protection authorities, and the data subjects should be provided with effective and enforceable rights and effective administrative and judicial redress."

In sum, the Commission ought to balance all of these elements and compare the level of data protection in the third country to the one in Europe.
Transfers by the way of BCR are specified in Article 47 GDPR. BCR have to fulfill certain criteria in order to make a data transfer to a third country lawful. They have to ensure all essential principles and enforceable rights of the GDPR to be considered an appropriate safeguard for third country transfers. Their purpose is to enable corporate groups to transfer data to entities within the same corporate group (Recital 108 GDPR). BCR will be approved by the Commission if they fulfill Article 47's criteria. BCR have indeed been generally accepted by Article 26 Par. 2 DPD as adequate safeguards, however they have not yet been mentioned explicitly in the DPD. [289]
Having regard to cloud computing, BCR shall only be approved by the Commission if they categorically have a binding character:

> Article 47 Par. 1 (a):
>
> "BCR are legally binding and apply to and are enforced by every member within the controllers' group of undertakings and those external subcontractors that are covered by the scope of the binding corporate rules, and include their employees"

[289] *Hullen*, in: v. d. Bussche/Voigt, Konzerndatenschutz, Teil 8, recital 55.

Moreover, data protection certifications and seals (Article 42 Par. 2 GDPR) can be used by the controller to provide evidence of a processor's compliance with the GDPR in the case of processing on their behalf (see 2.3.2.4). Moreover, the certification can provide evidence for appropriate safeguards concerning the level of data protection in order to permit a transfer to a third country. Consequently, the data protection certification can be important for both steps needed to render a third-country transfer lawful.

Appropriate safeguards can also be provided by means of standard data protection clauses by the Commission or standard data protection clauses adopted by a supervisory authority. In each model the contract has to be concluded between the controller transferring the data and the party receiving the data in the third country. Thus, the receiving party shall be bound to European data protection principles. According to Article 13 GDPR, although the person whose data is being processed is not part of the contract, this person has to be provided with information. Whereas standard contract clauses which are acknowledged by the EU commission (Article 46 Par. 2 (d)) have to be adopted by a supervisory authority and the Commission pursuant to the examination procedure referred to in Article 93 Par. 2, individual contract clauses of a controller need prior authorization from the competent supervisory authority (not the Commission), Article 46 Par. 3 (a) GDPR. Standard contract clauses are already acknowledged in Article 26 Par. 4 DPD and shall remain valid according to Article 45 Par. 9 GDPR.

Standard data protection clauses according to Article 46 Par. 2 (c) shall only be accepted by national supervisory authorities or according to Article 60 Par. 1 GDPR. [290]

In general, international arrangements e.g. between the European Union and the United States (see regarding the EU-US-Privacy Shield 2.4.3.1) will not be affected by the GDPR. [291] Recital 102 states that:

> "This Regulation is without prejudice to international agreements concluded between the Union and third countries regulating the transfer of personal data including appropriate safeguards for the data subjects. Member States may conclude international agreements which involve the transfer of personal data to third countries or international organisations, as far as such agreements do not affect this Regulation or any other provisions of Union law and include an appropriate level for the fundamental rights of the data subjects."

In fact, the judgment of the ECJ declaring Safe Harbor to be invalid [292] affected the different proposals for a GDPR and lead to the adoption of certain new principles in the GDPR's final version (see above). However, according to Article 45 Par. 9 and Article 46 Par. 5 GDPR, decisions adopted by the Commission on the basis of Article 25 Par. 6 or Article 26 Par. 4 DPD shall remain in force, until amended, replaced or repealed by a Commission Decision.

Data transfers to a controller or processor within the USA will, therefore, still be possible if the receiving party is bound to an adequate level of protection.

If appropriate safeguards have not been taken to guarantee an adequate level of data protection, the transfer of personal data to a third country can only be carried out if Article 49 GDPR's requirements are met. Thus, either the data subject has to give his or her consent (causing the same problems as described above, see 2.4.2.3) to the transfer or one of the legal permissions in Article 49 (b) to (g) should be applicable. Those permissions are similar to Article 6 GDPR's legal permissions (see

[290] *Hullen*, in: v. d. Bussche/Voigt, Konzerndatenschutz, Teil 8, recital 55, referring to Article 57 GDPR of the Commission's and Council's proposals.

[291] *Nebel/Richter*, ZD 2012, 407 (412).

[292] ECJ, Judgment of 6th October 2015 in Case C-362/14 - Maximillian Schrems v Data Protection Commissioner and Digital Rights Ireland Ltd.

2.4.2.2) for processing personal data. Note that Article 49 takes effect on the second step (whether the transfer to a third country is lawful) and not on the first step (whether the processing, itself, is lawful).

2.4.4 Technical and Organizational Measures

2.4.4.1 Under the DPD

Appropriate technical and organizational measures have to be provided in order to avoid data leaks, data loss and illegal forms of personal data processing, Article 17 Par. 1 Data Protection Directive. The core security objectives are availability, confidentiality, and integrity; in addition, transparency, accountability and portability also have to be taken into account. [293] As the DPD does not specify exactly which measures have to be taken, data controllers are, to some extent (and depending upon the practice of national supervisory authorities), flexible to adopt the appropriate measures. Existing ISO/IEC standards can be adopted and applied by data processing entities to ensure providing appropriate technical and organizational measures. They can be used as a general guide for initiating and implementing the IT security management process. [294]

In order to achieve the goal of enhancing (or guaranteeing) the safety of personal data, one of the crucial elements is the isolation of every client's computing on every level of the cloud computing stack. In other words, computing processes have to be protected from other parties who want to access it, so that personal data is literally "safe".

Moreover, the IT-infrastructure (networks, IT-systems, applications) has to be secure, even including physical resources, like buildings and employees. [295] Providing availability of data means ensuring timely and reliable access to personal data. Integrity implies that data is authentic and has not been maliciously or accidentally altered during processing, storage or transmission. Thus, remote administration of a cloud platform should only take place via a secure communication channel. [296]

Article 17 DPD states that the measures taken have to protect the personal data against unauthorized disclosure or access. The state of the art has to be respected in order to assess which measures are appropriate.

One of the technical means to ensure confidentiality is encryption which can protect data against illegal access, disclosure or alteration during its storage and transfer. Encryption can anonymize data or at least pseudonymised data (which complies with the principle of data minimization and data reduction) and thus protect personal data against misuse, especially against attacks from the Internet. [297]

Hence, for PRACTICE, one important tool to ensure confidentiality is encryption.

2.4.4.2 Under the GDPR

The GDPR will change the specification of technical and organizational measures. Article 32 GDPR regulates the controller's and processor's duties, regarding the detailed measures to be taken. Nevertheless, the core principles set out in Article 32 are similar to those developed by the DPD.Article 32 Par. 1 GDPR:

[293] See also *Art. 29-Working Party*, Opinion 05/2012, WP 196, 14

[294] For a list of ISO standards with further explanation see *German Federal Office for Information Security Technology*, BSI-Standard 100-1 Information Security Management Systems, p. 8.

[295] *German Federal Office for Information Security Technology*, Safety Recommendation for Cloud Computing Providers, p. 28 ff.

[296] *Art. 29-Working Party*, Opinion 05/2012, WP 196, p. 14 f.

[297] *Hartung/Storm*, in: Hilber, Handbuch Cloud Computing, Teil 4 recital 117; furthermore, sentence 3 of the attachment to par. 9 BDSG explicitly mentions encryption as a possible technical and organizational measure.

> "Taking into account the state of the art, the costs of implementation and the nature, scope, context and purposes of processing as well as the risk of varying likelihood and severity for the rights and freedoms of natural persons, the controller and the processor shall 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**; (emphasis added)
>
> (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."

Consequently, encryption of personal data will be a very useful tool to accomplish the task of ensuring integrity and confidentiality, set by Article 32 GDPR, since it is now explicitly mentioned in lit. (a) as a technical and organisational measure. Furthermore, according to Par. 2:

> "In assessing the appropriate level of security account shall be taken in particular of the risks that are presented by processing, in particular from accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to personal data transmitted, stored or otherwise processed."

As mentioned above, encryption-technologies are developed to prevent unauthorized access to data. This shall be accomplished by having regard to the state of the art and the costs of their implementation (Article 30 Par. 1 GDPR).
Moreover, when deciding on the amount of administrative fines in case of violations of data protection according to Article 83 Par. 2 lit. (d) GDPR regard shall be given on the degree of responsibility of the controller or processor taking into account technical and organisational measures implemented by them.
Hence, it will be necessary to always use encryption that is considered as "State-of-the-Art". To demonstrate compliance with the requirements of Par. 1 an approved code of conduct or an approved certification mechanism may be used (see in detail 2.3.2.4). Additionally, according to Recital 77 GDPR the European Data Protection Board [298] may issue guidelines inter alia "on processing operations that are considered to be unlikely to result in a high risk to the rights and freedoms of natural persons and indicate what measures may be sufficient in such cases to address such risk."
Finally, Article 32 Par. 4 GDPR deals with obligations for controllers and processors regarding their employees:

> "The controller and processor shall take steps to ensure that any natural person acting under the authority of the controller or the processor who has access to personal data does not process them except on instructions from the controller, unless he or she is required to do so by Union or Member State law."

[298] A board composed of the heads of the supervisory authorities of the Member States and the European Data Protection Supervisor - similar to the Art. 29 Working Party Articles 68 GDPR.

Thus, cloud providers need to ensure by establishing technical and organisational measures such as access or transmission controls that their employees who may have access to personal data process this data only on the instructions of the cloud user (the controller). [299]
Moreover, according to Article 28 Par. 3 S. 2 lit. (b) GDPR the processor, in our case a cloud provider, shall ensure that persons authorised to process the personal data have committed themselves to confidentiality or are under an appropriate statutory obligation of confidentiality.
Developing technologies using state-of-the-art encryption to enable privacy-preserving cloud computation is the main goal of PRACTICE. Cloud users and providers will be able to take technical measures as demanded by Article 32 GDPR when using PRACTICE technologies.

2.5 Other reforms by the GDPR

2.5.1 Transfers or Disclosures not Authorised by Union law

Regarding the requirement of a controller or processor to disclose personal data to a third country, Article 48 GDPR stipulates that:

> “Any judgment of a court or tribunal and any decision of an administrative authority of a third country requiring a controller or processor to transfer or disclose personal data may only be recognised or enforceable in any manner if based on an international agreement, such as a mutual legal assistance treaty, in force between the requesting third country and the Union or a Member State, without prejudice to other grounds for transfer pursuant to this Chapter.”

This negative clause is aimed at activities of third countries obliging providers (data controllers, processors) to disclose personal data following the NSA scandals and revelations of Edward Snowden. To protect persons within the EU from having their personal data transferred to a third country based on a third country ruling which is not compliant with the European data protection law, Recital 115 states that:

> “Some third countries adopt laws, regulations and other legal acts which purport to directly regulate the processing activities of natural and legal persons under the jurisdiction of the Member States. This may include judgments of courts or tribunals or decisions of administrative authorities in third countries requiring a controller or processor to transfer or disclose personal data, and which are not based on an international agreement, such as a mutual legal assistance treaty, in force between the requesting third country and the Union or a Member State. (. . .) Transfers should only be allowed where the conditions of this Regulation for a transfer to third countries are met. This may be the case, inter alia, where disclosure is necessary for an important ground of public interest recognised in Union or Member State law to which the controller is subject.”

[299] More detailed provisions regarding this can be found in the Annex to the first sentence of Section 9 of the German data protection act BDSG. According to this, where personal data are processed the internal organisation of an enterprise has to be arranged in such a way that e.g. due to access control unauthorized persons cannot gain access to data processing systems with which personal data are processed or used, mostly by using physical barriers, moreover, by providing means to ensure that persons entitled to use a data processing system have access only to the data to which they have a right of access, and that personal data cannot be read, copied, modified or removed without authorization in the course of processing or use and after storage or to ensure that data collected for different purposes can be processed separately. See in detail *Plath*, in: Plath, BDSG, 2nd Ed. 2016, 9 BDSG, Recitals 21 ff.

Thus, the European data protection law can require data controllers and processors to break a third country's law in order to comply with Article 48 GDPR if there is no international agreement regarding this case or if the GDPR's requirements regarding the regular data transfer to third countries are not met (see above 2.4.3.2). Moreover, a recent ruling by the United States Court of Appeals for the Second Circuit decided that a warrant issued under Section 2703 of the Stored Communications Act (SCA) cannot oblige Microsoft as an American company to disclose data stored in servers outside the United States to the US government. [300] The warrant was specifically referring to an e-mail account hosted in Dublin, and, therefore, stored within the EU. According to the decision, Section 2703 of the SCA "does not authorize courts to issue and enforce against US-based service providers warrants for the seizure of customer e-mail content that is stored exclusively on foreign servers."

2.5.2 Privacy by Design and by Default

One of the main innovations of the proposed Regulation refers to Privacy by Design, Article 23. The Privacy by Design principle requires all producers, data controllers, etc., to respect data protection issues whilst developing or implementing new IT-systems or products. [301]
Thus, any privacy issues shall already be addressed during the development of new technologies in order to find solutions from scratch. Data Protection by Design must particularly take into account the entire lifecycle management of personal data, from collection to processing to deletion, systematically focusing on comprehensive procedural safeguards regarding accuracy, confidentiality, integrity, physical security and deletion of personal data. Article 25 Par. 1 GDPR obliges the controller to respect the principle of Privacy by Design by regulating that he or she "shall (...) implement appropriate technical and organisational measures, such as pseudonymisation, which are designed to implement data-protection principles, such as data minimisation, in an effective manner and to integrate the necessary safeguards into the processing". Moreover, Recital 78 GDPR states that "(s)uch measures could consist, inter alia, of minimising the processing of personal data, pseudonymising personal data as soon as possible, transparency with regard to the functions and processing of personal data, enabling the data subject to monitor the data processing, enabling the controller to create and improve security features." Thus, encryption as an instrument of pseudonymisation (and, according to the approach described in 2.2.3.6, in some cases also as an instrument of anonymisation) and data minimisation will be one of the technical measures a controller can use to comply with the principle of Privacy by Design. In contrast to the addressing of the controller, the developer of a data processing technology is not directly addressed by Article 25 GDPR, it only addresses the controller(s). However, Recital 78 GDPR encourages producers of products, services and applications that process personal data to take into account the right to data protection when developing and designing such products, services and applications and to make sure that controllers and processors are able to fulfill their data protection obligations. Such an encouragement is no obligation for the developers to comply with the principles of Privacy by Design, however, it could apply significant pressure on developers to only develop technologies that fulfill the requirements of Article 25 Par. 1 GDPR, because otherwise controllers will not use their services anymore to avoid violations of the requirements of the GDPR. [302]
Nevertheless, a controller is often not able to influence the development of a technology. In particular, a cloud user in the role of the responsible data controller is not developing the technology the cloud provider is using. Even the cloud provider may just present their service based on technologies

[300] *United States Court of Appeals for the Second Circuit*, decision of 14 July 2016, Case 14-2985 - Microsoft v. United States, available at: `http://pdfserver.amlaw.com/nlj/microsoft_ca2_20160714.pdf`.

[301] *Decker*, Die neue europäische Datenschutzgrundverordnung - welche änderungen sind für deutsche Unternehmen zu erwarten?; *Schaar*, Privacy by design; Krempl, EU-Datenschützer fordert Einbau von Datenschutz in die Technik.

[302] *Plath*, in: Plath, BDSG, 2nd Ed. 2016, Art. 25 DSGVO, Recitals 7.

offered by third parties (such as software developers etc.). [303]

2.5.3 'Right to erasure' ("Right to be Forgotten")

One of the major innovations of the GDPR is the right to erasure, regulated in Article 17 GDPR, which takes up the so-called "right to be forgotten" established by the ECJ in its Google Spain decision in 2014 for search engines to remove links to webpages that appear when searching a person's name. [304] The GDPR now expands this right to all data controllers.

According to Article 17 Par. 1 GDPR personal data concerning the data subject shall be erased "without undue delay" by the controller if one of the grounds stated in Par. 1 lit. (a) to (f) applies, e.g. if the the personal data are no longer necessary in relation to the purposes for which they were collected or otherwise processed or if the data subject withdraws consent and where there is no other legal ground for the processing or if the personal data have been unlawfully processed.

Furthermore, to strengthen the right to be forgotten in the online environment (see Recital 66 GDPR), Par. 2 regulates that controllers who made the data public are obliged to take all reasonable steps to inform other controllers which are processing the data that the data subject has requested erasure of. The erasure shall include any links to, or copy or replication of those personal data. The test of reasonability refers to available technology as well as the cost of implementation.

However, no data shall be erased in accordance with Par. 3, for example, when processing of the personal data is necessary (a) for exercising the right of freedom of expression and information [305], or (b) for compliance with a legal obligation which requires processing of personal data by Union or Member State law to which the controller is subject, or for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller, or for (c) reasons of public interest in the area of public health or (d) for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes or (e) for the establishment, exercise or defence of legal claim.

However, a lot of details still remain unresolved. For example, the balance between the right of the public to be informed by archives and historical information, and the right of the individual to have the data deleted. This clash between the fundamental right to privacy and the right of the public to be informed could have been dealt with efficiently in the ECJ's recent Google Spain decision. Unfortunately, the ECJ stated very briefly that as a general rule, the data subject's rights override the interest of the public. [306] However, the ECJ also conceded that depending on the sensitivity of information stored, there are specific cases in which the interest of the person whose data is being processed may be outweighed by the public's interest in accessing that information; for instance, in cases of persons of public interest. For cloud computing it is, after all, important that the cloud user is able to compel the cloud provider to delete personal data. Hence, if the provider is processing data on behalf of the user, then the provider should be bound by such a contractual obligation to delete data - within the general contract framework needed for "order processing" (see 2.3.2).

[303] *Roßnagel/Richter/Nebel*, ZD 2013, 103 (105).

[304] *ECJ*, Judgment of 13 May 2014, Case C-131/12 - Google Spain SL/Google Inc. v AEPD/Mario Costeja Gonzalez; see regarding the "right to be forgotten" *Mantelero*, Computer Law & Security Review 2013, 229 ff.; *Tamó/George*, JIPITEC 2014, 71 ff.; *Spiecker gen. Döhmann*, Common Market Law Review 2015, 1033 (1038 ff.).

[305] See regarding the clash of the right to be forgotten and of freedom of expression on the Internet *Fazlioglu*, International Data Privacy Law, 2013, 149 ff.

[306] *ECJ*, Judgment of 13 May 2014, Case C-131/12 - Google Spain SL/Google Inc. v AEPD/Mario Costeja Gonzalez, Recital 81.

2.5.4 Significant Increase of Fines

The lack of enforcement is one of the most important concerns of the current data protection legislation. One of the actions taken by the GDPR to overcome these deficits is an increase of fines in Article 83 Par. 5 to a maximum of 4% of the global annual turnover of an infringing company or up to 20.000.000 EUR - which is comparable only to antitrust fines. The fines shall in each individual case be effective, proportionate and dissuasive. Moreover, according to Recital 150 S. 3 GDPR "(w)here administrative fines are imposed on an undertaking, an undertaking should be understood to be an undertaking in accordance with Articles 101 and 102 TFEU for those purposes." This could lead to the interpretation that the legal concept of an "undertaking" has to interpreted in an antitrust point of view. Thus, fines could not only be imposed against the (subsidiary) company acting unlawfully, but also against the parent company, which would lead to the result that the fine would not just cover 4% of the company violating the GDPR, but also 4% of the global turnover of the whole group. [307]

2.5.5 Notification of Personal Data Breaches

In case of personal data breaches, according to Article 33 Par. 1 GDPR the controller has to inform the supervisory authority without undue delay ("and, where feasible, not later than 72 hours after having become aware of it"), unless the personal data breach is unlikely to result in a risk to the rights and freedoms of natural persons. Moreover, the processor of data has to inform the controller about any personal data breaches without undue delay. The obligation to report data breaches covers all kinds of personal data. Even unauthorized access to data within the controller's company or agency is considered to be a data breach, and thus has to be notified to the supervisory authority.

Article 33 Par. 3 GDPR lists the minimum requirements that a notification has to meet. The notification has to include the approximate number of data subjects and data records concerned. It might be hard to tell how many data subjects or data records have been lost if a server that is used for cloud computing has been compromised, due to the scalability of cloud computing and the fast transfer of data sets. [308]

According to Article 34 Par. 1 GDPR, the controller shall communicate the personal data breach also to the data subject without undue delay when the personal data breach is likely to result in a high risk to the rights and freedoms of natural persons. However, Par. 3 lit. (a) stipulates that the communication to the data subject shall not be required if "the controller has implemented appropriate technological and organisational protection measures, and those measures were applied to the personal data affected by the personal data breach, in particular those that render the data unintelligible to any person who is not authorised to access it, such as encryption." However, since the time-period of the data breach is relevant for the assessment of the security of the encryption technology used, and the standard for the safe encryption continuously changes, regulations concerning the question of which level of technical security is adequate can unfortunately not be found in the GDPR's wording. [309] The Data Protection Authority is obligated to keep a public register of the types of breaches notified. Article 33 GDPR refers to all kind of data breaches, making no difference between third-party attacks (hacker etc.), employees, etc.

Still unresolved - and implicitly left to the Member States - is the issue of civil liability for data breaches, particularly if omitted breach notifications may constitute grounds for civil action.

Thus, applying state-of-the-art encryption technologies can exempt the controller to communicate a personal data breach to the data subject.

[307] *Faust/Spittka/Wybitul*, ZD 2016, 120 (123 f.).

[308] For the same reason it might be hard to gain informed consent for processing data in a cloud - see 2.4.2.3 and 2.4.2.4.

[309] *Marschall*, DuD 2015, 183 (189).

2.5.6 Right to Data Portability

The GDPR introduces in Article 20 the new right to data portability [310] , according to which "(t)he data subject shall have the right to receive the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and machine-readable format and have the right to transmit those data to another controller without hindrance from the controller to which the personal data have been provided". This right shall apply to the data subject if the processing is based on consent on a contract and if the processing is carried out by automated means. Moreover, according to Par. 2, where technically feasible, the data subject shall have the right to have the personal data transmitted directly from one controller to another. However, according to Par. 4 the right to data portability shall not adversely affect the rights and freedoms of others - which can especially be a problem when applying this right regarding personal data processed in social networks. Thus, cloud providers will have to install a technical environment with which they can provide the data subjects with their personal data if obtained, furthermore, they will have to be able to transfer the data to other controllers.

2.5.7 One-Stop-Shop

Another innovation of the GDPR is the introduction of a 'one-stop-shop' for the European data protection supervisory authorities. [311] According to Article 56 Par. 1 GDPR, the supervisory authority of the main establishment or of the single establishment of the controller or processor shall be competent to act as lead supervisory authority for the cross-border processing carried out by that controller or processor. Although the uniform interpretation and enforcement of data protection regulation was one of the main goals of the GDPR, in fact, the new regulation does not provide a consistent and clear "one-stop-shop" solution for all data protection issues. Instead, several exemptions within the GDPR can lead to the involvement of additional data protection authorities in other Member States. [312] Article 4 No. 16 GDPR defines "main establishment" regarding the controller as "the place of its central administration in the Union, unless the decisions on the purposes and means of the processing of personal data are taken in another establishment of the controller in the Union and the latter establishment has the power to have such decisions implemented, in which case the establishment having taken such decisions is to be considered to be the main establishment." Regarding the processor, it means the place of its central administration in the Union or, if the processor has no central administration in the Union, the establishment of the processor in the Union where the main processing activities in the context of the activities of an establishment of the processor take place. "Cross-border processing" means according to Article 4 No. 23 GDPR either "processing of personal data which takes place in the context of the activities of establishments in more than one Member State of a controller or processor in the Union where the controller or processor is established in more than one Member State" or "processing of personal data which takes place in the context of the activities of a single establishment of a controller or processor in the Union but which substantially affects or is likely to substantially affect data subjects in more than one Member State." Hence, under the GDPR, companies with establishments in several Member States will (only) have to deal with the supervisory authority of the Member State in which their main establishment is seated.

[310] See for more details *Zanfir*, International Data Privacy Law 2012, 149 (157 ff.); *Jülicher/Rttgen/Schönfeld*, ZD 2016, 358 ff.

[311] C.f. *Hullen*, in: v. d. Bussche/Voigt, Konzerndatenschutz, Teil 8, recital 39 ff.

[312] See for more details *Gierschmann*, ZD 2016, 51 ff.

2.5.8 Data Protection Impact Assessment

Another innovation of the GDPR is the data protection impact assessment, regulated in Article 35 GDPR and considered in Recitals 76 and 84 of the GDPR. According to Article 35 Par. 1 GDPR, the impact assessment refers to the fact that some types of processing personal data, in particular when using new technologies, are likely to result in a high risk to the rights and freedoms of natural persons. Therefore, the controller shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data. In order to full the impact assessment, Article 35 mentions three exampled cases, when the impact assessment shall in particular be required (Par. 3) and states the minimum assessment requirements for the controller to act lawfully (Par. 7).

In addition to that, Recital 84 underlines that an impact assessment is implemented to promote the compliance with the GDPR [313] as the evaluation resulting from the origin, nature, particularity and severity of risks should be taken into account when determining the appropriate measures. To avoid legal uncertainty as the abstract legal concept of "high risk" gives room for interpretation, it is recommended that the controller undertakes the assessment, when there is a reasonable chance that a risk to the rights and freedoms of natural persons is given.

[313] *Hansen*, DuD 2016, 587 (588).

Chapter 3

Legal Case Studies

3.1 Encrypted Databases - Encrypted HANA

3.1.1 Functioning

Stealing private information by collecting data is a significant problem, especially for online applications. One of the solutions is to encrypt sensitive data, in order to make sure that no unauthorized person is able to use the sensitive data. However, if personal data is encrypted, some applications and programs may not be able to handle and further process that encrypted data.
However, Encrypted HANA is based on CryptDB and is able to solve this problem. Encrypted HANA addresses two threats: The first refers to a "curious" database administrator who tries to learn and, in the worst case, spy on personal data by snooping on the database management system (DBMS) [1] , see Figure 1. The second threat concerns an external attacker who gains complete control of an application and database management system. [2] To avoid these two threats, Encrypted HANA minimizes the amount of confidential information revealed to the database management system server whilst providing a variety of queries over the encrypted data, thus enabling further processing of data.

3.1.1.1 Three Main Ideas of Encrypted HANA

Encrypted HANA tries to solve this problem by using three main ideas: [3]

Execution of SQL- Queries Over Encrypted Data

The main idea of Encrypted HANA is based upon an SQL-aware encryption strategy. [4] SQL-queries consist of a well-defined set of primitive operators, such as equality checks, order comparisons, aggregates, and joins. By adapting known encryption schemes and using new privacy-preserving cryptographic method for joins, HANA encrypts each data item in a way that allows the DBMS to compute the transformed data. Encrypted HANA is efficient because it mostly uses symmetric-key encryption, avoids fully homomorphic encryption, and runs on unmodified DBMS software (by using user-defined functions).

[1] *Popa/Redfield/Zeldovich/Balakrishman*, CryptDB: Protection Confidentiality with Encrypted Query Processing, p. 1 f.

[2] *Popa/Redfield/Zeldovich/Balakrishman*, CryptDB: Protection Confidentiality with Encrypted Query Processing, p. 2 f.

[3] *Popa/Redfield/Zeldovich/Balakrishman*, CryptDB: Protection Confidentiality with Encrypted Query Processing, p. 1 ff.

[4] *Popa/Redfield/Zeldovich/Balakrishman*, CryptDB: Protection Confidentiality with Encrypted Query Processing, p.1 f.; Popa /Zeldovich/Balakrishman CryptDB: A Practical Encrypted Relational DBMS, p. 3.

Adjustable Query-Based Encryption

The encodings differ in their encryption and safety. [5] Some methods of encryption are more easily decrypted than others but have to be used for certain queries over the information stored on the DBMS. To prevent all possible disclosures of the encrypted data, Encrypted HANA changes the encryption scheme to some specific data elements of the SQL-queries, depending on the queries observed at run-time. For the efficient implementation of these adjustments, Encrypted HANA uses multiple layers of encryption.

Chain Encryption Keys to User Passwords

One of the principle ideas of Encrypted HANA refers to chaining the encryption. [6] With this method, each data item on the database proxy server can be decrypted only via a chain of keys rooted in the password of one of the users with access to that specific data. If the user is not logged into the application and if the administrator or an external attacker does not know the password, the description of the data cannot be overruled. To create that chain of keys, Encrypted HANA allows the developer to provide policy annotations to the application's SQL schema, specifying which users have access.

3.1.1.2 Benefits from Encrypted HANA

Encrypted HANA ensures that the sensitive data is never available in plaintext at the DBMS (Database Management System) server. The information sent to the DBMS depends on the classes of computation required by the application's queries; thus, the DBMS cannot compute the encrypted results that involve computation classes not requested by the application.

3.1.1.3 Encrypted HANA's Architecture

Encrypted HANA's architecture consists of two parts: a database proxy and an unmodified Database Management System server. Encrypted HANA uses user-defined functions to perform cryptographic operations in the database management. [7]

For many years, the problem in using encryption was that the programs could not handle strongly encrypted files. Encrypted HANA avoids these problems by intercepting all SQL queries in a database proxy which rewrites queries to execute on encrypted data. The system allows the users to send queries to an encrypted set of data and get the answer they need from it without ever self-decrypting the stored information. [8] The database proxy encrypts and decrypts all data, and changes some query operators while preserving the semantics of the query. The unmodified DBMS never receives decryption keys of the plaintext; thus, the administrator or an external attacker never sees sensitive data. Hence, no one can access the private data without authorization (threat 1). [9]

To protect and shield against the application, the database proxy and the unmodified DBMS would

[5] *Popa/Redfield/Zeldovich/Balakrishman*, CryptDB: Protection Confidentiality with Encrypted Query Processing, p. 2.

[6] *Popa/Redfield/Zeldovich/Balakrishman*, CryptDB: Protection Confidentiality with Encrypted Query Processing, p. 2.

[7] *Popa/Redfield/Zeldovich/Balakrishman*, CryptDB: Protection Confidentiality with Encrypted Query Processing, p. 3.

[8] *Popa/Redfield/Zeldovich/Balakrishman*, CryptDB: Protection Confidentiality with Encrypted Query Processing, p. 2.

[9] *Popa/Redfield/Zeldovich/Balakrishman*, CryptDB: Protection Confidentiality with Encrypted Query Processing, p. 2.

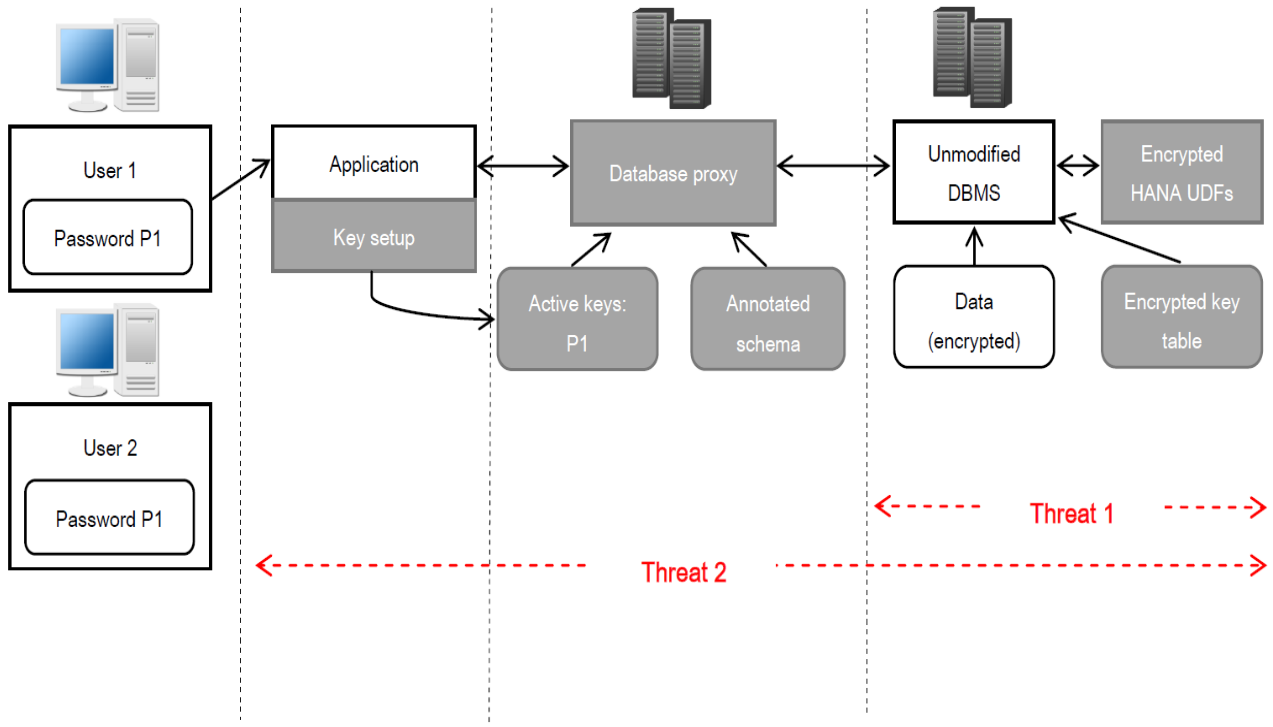


Figure 3.1: Architecture of Encrypted HANA

enter into an agreement (threat 2). [10] The developers annotate their SQL schema to define different principals whose keys will allow access to decrypt the different parts of the database. They also make a small change to their applications to provide encryption keys to the proxy. Then the database proxy determines which parts of the database should be encrypted with which key. As a result, Encrypted HANA can guarantee the confidentiality of the data belonging to users who are not logged in the database proxy server during the compromise and who do not log in until the compromise is detected and fixed by the administrator. [11]

Encrypted HANA can protect data confidentiality, however, it cannot ensure integrity or completeness of results. Therefore, it is possible that a malicious administrator or external attacker may gain access to the application, the database proxy or the DBMS, and delete the existing data.

3.1.1.4 Queries over Encrypted Data

As mentioned already, Encrypted HANA guarantees security of personal data. It enables the execution of SQL queries on encrypted data without any need to change the existing applications to work with Encrypted HANA. The DBMS's query plan for an encrypted query is exactly the same as for the original query. Only the operators comprising the query, such as selections, projections, joins, aggregates and orderings, are performed in ciphertexts and use modified operators in some cases. Encrypted HANA proxy accumulates a secret master key (key), the database scheme, and the current encryption layers of all columns. The DBMS only gets access to an anonymized scheme, encrypted user data, and some auxiliary tables used by Encrypted HANA. Encrypted HANA also supplies the

[10] *Popa/Redfield/Zeldovich/Balakrishman*, CryptDB: Protection Confidentiality with Encrypted Query Processing, p. 2.

[11] *Popa/Redfield/Zeldovich/Balakrishman*, CryptDB: Protection Confidentiality with Encrypted Query Processing, p. 2.

server with Encrypted HANA-specific, user-defined functions that enable the server to calculate on ciphertexts for certain operations. [12] Processing a query in Encrypted HANA takes four steps:

1. At first the applications distribute a query, which the proxy obstructs and rewrites: it anonymizes each table and column name, and, by using the master key, encrypts each constant in the query with an encryption scheme appropriate for the required operation [13]

2. Subsequently the proxy examine if the DBMS should be given keys to adjust encryption layers before carrying out the query, and if so, issues an UPDATE query at the DBMS that summon a UDF to adjust the encryption layer of the appropriate columns [14]

3. In the third step, the proxy returns the encrypted query to the DBMS which it carryies out using standard SQL [15]

4. In the end, the DBMS sends the encrypted query result back, which the proxy decrypts and delivers to the application. [16]

3.1.1.5 End User Applications with CryptDB as an Underlying Technology

Google recently deployed a system for performing SQL-like queries over an encrypted database following the CryptDB design. [17] Their service is able to use the encryption building blocks from CryptDB, rewrite queries and annotate the schema, as in CryptDB. [18] Lincoln Labs also started working with CryptDB and added its design on top for their D4M Accumulono-SQL engine. [19] Moreover, SAP SE developed a privacy preserving system based on CryptDB called SEEED. Based on HANA database their system SEEED uses most of the building blocks as well as the adjustable encryption (onion) strategy. [20] In addition to that, SEEED and HANA are used as main components of the PRACTICE prototype in the field of secure supply chain management in the aerospace industry. The prototype combines machine learning and order-preserving encryption to ensure privacy-preserving forecasting of maintenance demand based on condition data (see 4.1 of Deliverable D 24.4).

3.1.2 Legal Evaluation and Risk Assessment

3.1.2.1 Introduction: Legal Classification of the Involved Parties and the Data Processing Activities

There are two steps needed to run Encrypted HANA. Before queries can be run over encrypted data on the DBMS-server, the data has to be stored on the server. If we assume that the original (plaintext) data provided by the data subject (the affected person) is personal data, then the storage on the server

[12] *Popa/Redfield/Zeldovich/Balakrishman*, CryptDB: Protection Confidentiality with Encrypted Query Processing, p. 4.

[13] *Popa/Redfield/Zeldovich/Balakrishman*, CryptDB: Protection Confidentiality with Encrypted Query Processing, p. 4.

[14] *Popa/Redfield/Zeldovich/Balakrishman*, CryptDB: Protection Confidentiality with Encrypted Query Processing, p. 4.

[15] *Popa/Redfield/Zeldovich/Balakrishman*, CryptDB: Protection Confidentiality with Encrypted Query Processing, p. 4.

[16] *Popa/Redfield/Zeldovich/Balakrishman*, CryptDB: Protection Confidentiality with Encrypted Query Processing, p. 4.

[17] See `http://css.csail.mit.edu/cryptdb/`.

[18] *Popa*, Research Statement, p. 3.

[19] *Popa*, Research Statement, p. 3.

[20] See `http://css.csail.mit.edu/cryptdb/`.

has to be qualified as "processing," according to the DPD. During this first step, the entity which transfers the data is then simultaneously qualified as the controller as well as the processor. The storage of data is expressly mentioned in Article 2 (b) of the DPD as an action considered processing. During the second step, when the data has already been stored on the DBMS-server, the DPD will cover the computation of the data, itself. According to Article 2 (b) of the DPD, running the queries means using, aligning or combining, and consulting the data. The controller would still be the client and the processor would now be considered to be the provider of Encrypted HANA, as they would do the actual computation.

Therefore, we will have to carefully assess in the following whether the DPD is applicable, and in particular if personal data is really being affected or sufficiently anonymized by means of encryption:

3.1.2.2 Applicability of the DPD

As the data which is about to be transferred to the DBMS-server will be encrypted and only stored as ciphertext, it may fall out of the scope of the DPD, as it is no longer "personal data" from the perspective of DBMS, etc. As shown above, the effect encryption has on personal data is considered controversial.

According to the *absolute approach*, an anonymization and, therefore, the elimination of the data's connection to a data subject is only achieved when absolutely no one is able to decrypt the data. The client using Encrypted HANA is able to use the key and decrypt the data stored on the DBMS. Hence, regardless of the encryption, personal data is processed by the DBMS provider on behalf of the user of Encrypted HANA and, therefore, the Directive would be applicable. The absolute approach does not distinguish between those who are able to decipher the data and those who are not. Thus, even though the encrypted data is not readable for the controller (without considerable effort) it is still to be considered personal data as, at least theoretically, one person could access the personal data.

In contrast, the relative approach distinguishes between persons able to decrypt the cipher text (using reasonable efforts) and those who are not. The client using Encrypted HANA to run queries over their data is able to decipher the data they want to store on the DBMSserver. [21] The provider running the DBMSserver, on the other hand, is thus unable to decrypt the data. Moreover, if data is transferred to an entity unable to relate the data to certain persons (i.e. if the encryption standard is sufficient), the Directive is not applicable on this transfer - as due to the encryption, personal data is not affected any more. In other words, the applicability of the data protection law depends on the party receiving the data, not the one sending it. [22] It is not important if the party giving away the data is able to relate the data to a data subject as long as the receiving party is unable to.

Furthermore, the legal assessment depends on the standards required for encryption. Following the approach that a state-of-the-art encryption is sufficient, the encryption used by Encrypted HANA should be adequate. If an absolute encryption is demanded, Encrypted HANA's level of encryption might not be considered to be sufficient, especially concerning the insecure ways of encryption referred to in 3.1.1.1.2 of this chapter.

Assuming that state-of-the-art-encryption is seen as sufficient, since the Encrypted HANA-provider would not be able to decrypt the data, the encrypted data would not be qualified as personal data (following the relative approach). From the perspective of the absolute approach, there would be at least one entity able to decipher the encrypted data - the Encrypted HANA client - thus the data remains to be personal data (for everyone).

[21] Even if the single employee working with Encrypted HANA might not be able to decrypt all data, the legal entity he is working for, is considered to be the controller. The legal entity has in the end access to the data they work with in plaintext.

[22] *Dammann*, in: Simitis, BDSG, par. 3, recital 34.

3.1.2.3 Compliance with Existing and Future Data Protection Law

3.1.2.3.1 Compliance with the DPD

As mentioned above, the Directive can be applicable to the storage and computation on encrypted data with Encrypted HANA depending on the approach that has been chosen in relation to the notion of personal data and the level of encryption. As neither the DPD nor the GDPR implement a clear definition of personal data, the rather relative approach of the ECJ might answer the dispute whether the Directive is applicable on the processing of encrypted data with Encrypted HANA or not (see the ECJs approach, 1.2.1.4.1). However, we have to do the analysis using a two-fold approach in order to take into account a worst-case-scenario (if, contrary to our legal opinion, the absolute approach will prevail in the future):

If in spite of encryption, we assume the applicability of the DPD, the principle of prohibition of data processing without explicit consent or legal permission would come into effect.

If the data subject was sufficiently informed about the computation carried out and had freely given their consent, the controller in effect complied with the requirements of the Directive. However, we have to note that the required information concerns every kind of processing data, particularly the purposes, etc.

If consent is not available or not given, the processing may only take place in the case of explicit legal permissions such as outweighing interests of the processor or fulfilling contractual obligations, etc. (in the relationship with the data subject - not in relationship with the cloud provider and cloud user). However, this could hardly have been foreseen for all users of Encrypted HANA in every individual case.

Hence, even given the problems of providing sufficient information, it is highly recommended to obtain the data subject's consent for the storage of personal data on Encrypted HANA and the computation that runs with it. With regard to the sufficient information, it is advisable to use only physical machines in certain locations (within the EEA) as DBMS-servers, so the data subject can be told exactly where his data will be stored - thus, the problem of sufficient information can be minimized.

Another option to comply with the Directive regarding the transfer of the data to the DBMS-server would be a contractual framework that binds the Encrypted HANA provider legally to the user so that the provider would process the personal data on behalf of the controller in accordance with Article 17 of the DPD "order processing". Thus, the controller would be treated as if they were running the Encrypted HANA-technology themselves. Therefore, there would be no 'transfer' to another entity when the data is stored on the DBMS-server and, no permission would be needed to do so. On the other hand, to run queries over the data, they would still need permission (see 2.4.3.1).

If the DBMS-server is located in a state outside the EU/EEA, then the transfer of data is only permitted either if this state has an adequate level of protection or if adequate safeguards had been adduced, as mentioned above (2.4.3.1). Even if the processing would be carried out on behalf of the controller, as described above ("order processing"), these requirements have to be met by the controller.

Reengineering the query-results in order to obtain personal information should be as difficult as possible. If not, providing a query-result would lead to transferring personal data to the entity receiving it - thus bringing the DPD into play again.

The Encrypted HANA provider additionally has to comply with the Directive's requirements for technical and organizational measures to ensure data safety. Unauthorized access to the data has to be prevented.

By making computation of encrypted data possible, Encrypted HANA effectively minimizes the amount of personal data that has to be processed. [23]

[23] *Schaar*, Privacy by Design.

3.1.2.3.2 Compliance with the GDPR

Encrypted Data as not Being Qualified as Personal Data The application of the GDPR - just like the DPD - depends upon personal data being processed. As already mentioned, the forthcoming Regulation does not, unfortunately, solve the dispute concerning the approach for defining 'identifiable' data. However, recital 23 of the proposals for a Regulation states that all means reasonably *likely* to be used by the controller or by any other person should be taken into account. Whilst the words 'by any other person' might suggest an absolute approach, it is crucial to comprehend that only the means reasonably likely to be used have to be taken into account.

Concerning the specific case of Encrypted HANA, it is not reasonably likely that this state-of-the-art encryption could be overcome with reasonable efforts. Hence, the storage and computation of the encrypted data on the DBMS-server will not be affected by the GDPR, since no personal data will be processed by the Encrypted HANA-provider. It is important to understand that the relative approach (as it is followed here) does not treat the data processed by the DBMS-server provider as "encrypted data" for him as defined by Article 4 Par. 2b of the LIBE-proposal. The DBMS-server provider is not able to identify the affected persons using the encrypted data; therefore, for him the data is not encrypted personal data because it is not personal data, at all (from his perspective). However, from the perspective of those who are able to decrypt the data (the Encrypted HANA user), we have to qualify the data as personal data.

This distinction between different parties processing the data is the main disparity between the relative and the absolute approaches. *Technically*, whilst the DBMS-server provider stores encrypted data, they do not process "encrypted data" according to the relative approaches - since there is no personal data anymore.

However, as described above (see 2.2.3) we cannot exclude that in the political process of adopting the GDPR an absolute approach may gain approval. For this worst case scenario the options to comply with the GDPR when using Encrypted HANA shall be described.

Absolute Approach: Encrypted Data as Personal Data The provider of the DBMS may be considered as a processor on behalf of the user of Encrypted HANA, whereas the user's client (client of the controller) would be the affected person, the data subject. In order to comply with the GDPR, the transfer to the provider and the processing done by them should be constituted as 'order processing', as described in 2.3.2. The controller (the Encrypted HANA user) can ensure that appropriate technical safeguards as described in 2.4.4.2 have been taken if the provider is offering an encrypted database, like Encrypted HANA - fulfilling then their duties following Article 22 GDPR (see 2.3.2.1). To establish an order processing compliant with the GDPR, a contract between the provider and the user has to be concluded subduing the provider to the user's (controllers) instructions. It should enable the controller to document the processing as Article 28 GDPR demands and obliges the processor to take technical and organizational measures demanded by Article 30 GDPR (see 2.4.4.2). The controller has to report data breaches to the supervisory authority; hence, the contract with the processor should oblige him to inform the controller if such a breach occurs. If the controller has implemented appropriate technological and organizational protection measures such as a high level of encryption technique, according to Article 32 GDPR the controller does not need to inform the data subject about the breach (see 2.5.5).

A risk analysis of the potential impact of the data processing according to Article 32a GDPR (but only included in the LIBE-proposal) has to be carried out by the controller; moreover, according to Article 33 GDPR, under certain circumstances a data impact assessment should be carried out by either the controller or the processor. In this case, the contract regulating the order processing should clarify who will be responsible for this task.

In order to enable the user of Encrypted HANA (the controller) to comply with their duties, the provider the DBMS should apply for a certification (a privacy seal as described in 2.3.2.4) so that they can guarantee the processing in compliance with all technical and organizational measures required by the regulation. Such a certificate or privacy seal should ensure that all possible clients of the Encrypted HANA solution can comply with the Regulation's requirements to control the processor (the Encrypted HANA provider) processing on their behalf. If the provider has been certified the users of Encrypted HANA would not have to monitor and prove the compliance on their own, but rather be able to check the validity of the certificate (the data protection seal). For a DBMS-provider aiming at offering his service to multiple users, a certification is therefore highly recommended.
If the provider of the DBMS-server falls under the jurisdiction of a third country an order processing is not impossible (see 2.3.2.6); however, certain measures have to be taken in order to ensure adequate safeguards to comply with the GDPR. As described under 2.4.3.2, those measures may be an adequate acknowledgment of the EU Commission (Article 42 GDPR), 'binding corporate rules' (Article 42 Par. 2 a) GDPR), a European Data Protection seal (Article 42 Par. 2 aa) GDPR) (see 2.3.2.4), standard contract clauses (Article 42 Par. 2 c) GDPR) or contract clauses approved by a supervisory authority (Article 42 Par. 2 d) GDPR). If the controller is not able to ensure those measures, he has to obtain the data subject's consent or the processing has to take place on the basis of one of the permissions in Article 44 GDPR. The data protection seal can also be used in order to provide evidence for a lawful transfer to a third country.
Hence, the DBMS-server provider's chances of attaining a certification are even greater if they are based in a third county. According to Article 43a of the LIBE-proposal, the controller and the processor have to notify the supervisory authority and obtain prior authorization before disclosing personal data to a third country authority because of a third country's judgment, a court tribunal or a decision of an administrative authority (see 2.5.1).
Moreover, the principle of privacy by design will be explicitly included in the data protection act. In addition, data controllers have to continuously check and improve their systems if new challenges from the perspective of privacy (new risks, etc.) come to the light. Hence, if a new system using Encrypted HANA is established, privacy issues have to be dealt with when developing the new technology as well as their usage, from the implementation of its software on the client's system and the application server to the encryption and storage of data on the DBMS and the computation over this encrypted data. Not only the accuracy, confidentiality and integrity, but also the physical security of the system has to be kept in mind. It should be ensured that the client is able to use the system while giving away the least amount of personal data possible. All those requirements are directed by the encrypted database system, Encrypted HANA.
According to Article 14 GDPR, before the controller collects their clients' data, the controller will have to provide the clients with information. Therefore, among other information, the controller will have to inform the clients who the processor (the DBMS-server provider) will be, where the server will be located and for which purposes the controller intends to process the data (see 2.4.2.4). If the client exercises their rights according to Article 17 GDPR (the so called 'right to erasure' or 'right to be forgotten'), the user of Encrypted HANA will have to ensure the complete deletion of the clients data from the DBMS. According to Article 37 GDPR, the user of Encrypted HANA and the provider of the DBMS server shall designate a data protection officer.

3.2 Secret sharing

3.2.1 Sharemind

3.2.1.1 Functioning

Sharemind is a cloud-ready data analysis system for securely processing confidential information. [24] By its design, it provides security without the risk of an insider attack. The input data never leaves the hand of the owner, only final results of the computation are shared with partners. It provides privacy because private information can be processed without compromising the data subject's rights and convenience of use, Sharemind can be run in a cloud and is compatible with existing tools. [25] Sharemind is designed to be deployed as a distributed secure computation service that can be used for outsourcing data storage and computations by splitting personal or secret information between a minimum of three servers, to ensure the security of the data. [26] Sharemind Server is an application and a database server that uses secret sharing technology to store and process information without leaking it.

To achieve the best efficiency and privacy, three servers must be deployed by separate organizations that will likely not collude. The system is capable of performing computations on input data without compromising its privacy. Moreover, because the system is based on solid cryptographic foundations, it can process data whilst shielding data even from access of the server administrator.

3.2.1.1.1 Architecture of Sharemind

Sharemind uses Secure Multiparty Computation and secret sharing to protect the personal data of the user. [27]

3.2.1.1.2 Secure Multiparty Computation

Secure Multiparty Computation refers to a field of cryptography that deals with protocols involving two or more participants, who want to mutually compute a useful result. [28] Every party will provide an input value and learn only the result of their own individual value so that nobody is able to access all the information. [29] In the case of data aggregation algorithms, it is generally not possible to learn the inputs of other parties from the result. [30] Figure 3.2 shows the data storage process with three servers. The data is collected from the users or exists already on other servers and is sent to the three servers. A data donor distributes the data into parts/shares using secret-sharing and sends one random share of each value to a single server.

If the Sharemind technology is used to compare information from two entities in such a way that no one knows the others values, then both entities function as data donors. It can be necessary that one donor specifies what kind of information the other donor has to provide from its database, for example the IDs of the persons whose data is about to be compared).

[24] An overview of the project can be found under: `https://cyber.ee/en/security-technologies/sharemind/`.

[25] *Cybernetica*, Sharemind - Your secure service platform or data collection and analysis, p. 2.

[26] *Bogdanov*, Sharemind: programmable secure computations with practical applications, p. 30.

[27] *Kamm/Willemson*, Secure Floating-Point Arithmetic and Private Satellite Collision Analysis, p. 5.

[28] *Bogdanov*, Sharemind: programmable secure computations with practical applications, p. 24.

[29] *Kamm/Willemson*, Secure Floating-Point Arithmetic and Private Satellite Collision Analysis, p. 2.

[30] *Kamm/Willemson*, Secure Floating-Point Arithmetic and Private Satellite Collision Analysis, p. 2.

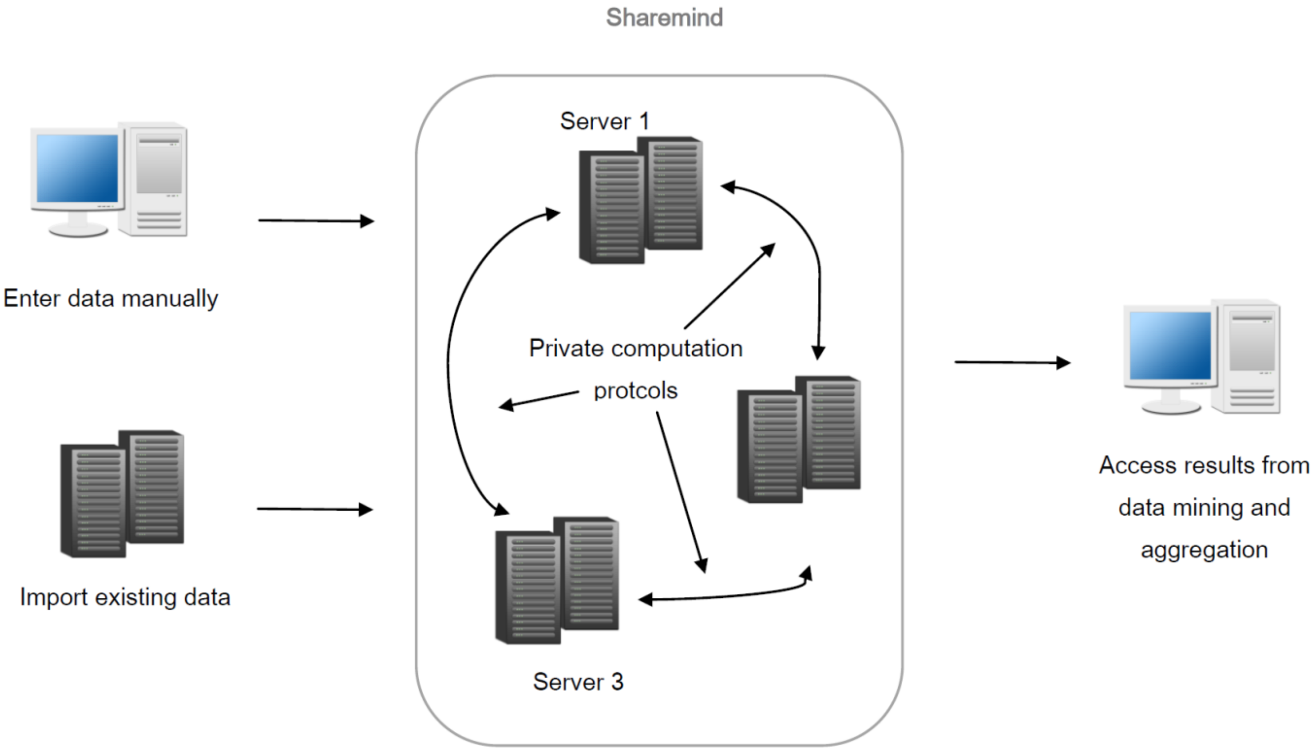


Figure 3.2: Function of Sharemind, three Sharemind Servers were deployed by three independent organizations, the information is divided between the three and every one of them receives a part of the information and works with it. At the end every organization sends his results back to the client.

The separation of servers between input donors and servers is useful as it does not force every party to run secure multiparty computation protocols. [31]
After the data has been transmitted and stored, the server can perform computations on the shared data; however, the server does not share the information with other servers. This is done so that none of them can reconstruct the input values. [32] The number of three servers is used for efficiency and is needed to guarantee the security during the computation; otherwise, it would be too easy to reconstruct the data. Moreover, an increase of servers reduces any risk of collusions. Secure multiparty computation protocols specify which messages the server should exchange in order to compute new shares of the value that corresponds to the result. [33] After finishing the computation, the results of the servers are transmitted and published to the client of the computation (Figure 1: the user). The servers send the share of the results to the user who reconstructs the real result. [34]

3.2.1.1.3 Secret-Sharing

In Sharemind, each party will receive one share of every secret value. The original secret can only be reconstructed by collecting all the shares of a value and adding them up using the addition operation in the ring. [35]

[31] *Kamm/Willemson*, Secure Floating-Point Arithmetic and Private Satellite Collision Analysis, p. 3.
[32] *Kamm/Willemson*, Secure Floating-Point Arithmetic and Private Satellite Collision Analysis, p. 3.
[33] *Kamm/Willemson*, Secure Floating-Point Arithmetic and Private Satellite Collision Analysis, p. 3.
[34] *Kamm/Willemson*, Secure Floating-Point Arithmetic and Private Satellite Collision Analysis, p. 3.
[35] *Bogdanov*, Sharemind: programmable secure computations with practical applications, p. 34.

3.2.1.1.4 Use Case: Secure Floating-Point Arithmetic and Private Satellite Collision Analysis and Sharing of Medical Data

One prominent case regarding the use of Sharemind refers to the delicate issue of sharing information about locations of satellites, etc., in order to avoid collisions. In the orbit, nearly 7,000 spacecraft are flying around the Earth. [36] In the year 2009, two communication satellites belonging to the USA and Russia collided in orbit because the nations did not talk about the trajectory of their satellites. [37] The two orbital planes intersected at a nearly complete right angle, resulting in a collision velocity of more than 11 km/s.[38] The locations and orbits of the communication satellites are sensitive information; hence, governments or private enterprises want to protect this data. [39] Whilst a trusted third party gathering all the data and performing analysis could be a solution, this party would still need disclosure of information of all the parties involved, thus endangering privacy and security of data. [40] By its design, it ensures secrecy of information by using Multiparty Computation and Secret Sharing Sharemind, and can, therefore, be used for calculating the probability of a collision between two satellites. [41] Using secure multiparty computation instead of a trusted third party could solve the problem of disclosure, and would be more practical. The satellite operators would choose three independent parties as the data host/servers. [42] Subsequently, the operators would secret-share their data and upload the shares to the three servers. At this point, collision analysis is a collaborative effort between the three hosting parties and the satellite operator, who can query the results of the analysis. [43]
The same method could potentially be used for many types of sensitive information. For instance, for private health data, in order to ensure that no unauthorized person is able to obtain health information.

3.2.1.1.5 Difference between Sharemind and Encrypted HANA

The two presented solutions in this report, Sharemind and Encrypted HANA, use different techniques to avoid handling personal data. Sharemind breaks each value down to several random fragments, so that the information is anonymized. Encrypted HANA works with encrypted queries and encrypted data so that the administrator or an external attacker does not have access to the personal information.

3.2.1.2 Legal Evaluation and Risk Assessment

3.2.1.2.1 A Legal Classification of the Involved Parties and the Data Processing Activities

Sharemind requires three steps: the donors have to be informed whose data shall be provided; the data has to be divided; and then stored on the different servers. If it is necessary for one data donor to specify whose information the other donor has to provide, this has to be considered as processing of personal data in a legal sense. It would then be inevitable to identify the data subjects whose information is needed for the purposes of computation. The transfer of the ID information would need the data subject's consent or an explicit legal permission. Alternatively, all data can be loaded to Sharemind and securely joined to them using ciphertexts. This would reduce the amount of personal

[36] *NASA*, NSSDC Master Catalog, last accessed February 16, 2014 available at: `http://nssdc.gsfc.nasa.gov/nmc/`.

[37] *NASA*, Orbital Debris Quarterly News, 2009, vol. 13, Issue 2, 1 f.

[38] *NASA*, Orbital Debris Quarterly News, 2009, vol. 13, Issue 2, 1 f.; Kamm/Willemson, Secure Floating-Point Arithmetic and Private Satellite Collision Analysis, p. 6.

[39] *Kamm/Willemson*, Secure Floating-Point Arithmetic and Private Satellite Collision Analysis, p. 8.

[40] *Kamm/Willemson*, Secure Floating-Point Arithmetic and Private Satellite Collision Analysis, p. 8.

[41] *NASA*, Orbital Debris Quarterly News 2009, Vol. 13, Issue 2, 1 f; Kamm/Willemson, Secure Floating-Point Arithmetic and Private Satellite Collision Analysis, p. 1.

[42] *Kamm/Willemson*, Secure Floating-Point Arithmetic and Private Satellite Collision Analysis, p. 8.

[43] *Kamm/Willemson*, Secure Floating-Point Arithmetic and Private Satellite Collision Analysis, p. 8.

data shared during the work done with Sharemind and, therefore, comply with the principle of Privacy by Design. [44]

Before the data is stored on the different servers, it has to be divided. This process must be carried out through personal data in plaintext. The problem regarding the applicability of the Directive is not the qualification of this data as personal data, under the terms of Article 2 (a) of the directive; it is, rather, whether dividing of personal data still has to be seen as processing personal data. In Article 2 (b) the directive mentions the alteration of data as processing. However, as described above, this refers to the alteration of content, not of its appearance. [45] The secret-sharing of personal data by dividing it does not fall under the Directive's scope.

Once the data has been divided, it will be stored on the different servers. If the data chunks were to be considered personal data (according to the absolute approach which is in principal not followed here, see 3.2.1.2.2), this kind of processing also would be qualified as processing of personal data. Even so, we have to state again that we consider this approach to be extending the scope of the DPD in an tremendous way.

Concerning Sharemind, it is not as easy to determine who the controller is: The role of the controller is not fixed to one of the many participants (at least two data donors, three data server providers and the user). The entity (user or in other terms "client") who is in charge will likely be the one who has initiated the research carried out with Sharemind and decided to use Sharemind. According to the relevant criteria regarding the definition of "controller" - in particular, who can determine the purposes of data processing, etc. - it shall be the user (the client) who initiates the process.

The provider of Sharemind can be one of the server providers who has the technical know-how to use the technology. However, this does not imply that they decide how the data processing is done with Sharemind. They should, rather, be seen as a usual provider of a cloud solution. In a way, Sharemind can thus be compared to Software as a Service (SaaS).

If more than one participant is determining the purpose of the processing, they each have to ensure compliance for all processing that is carried out.

Not easy to be answered (and still generally unresolved) is the issue of joint controllership if the "controllers" (the users, the clients) have initiated the data processing in such a manner that they have acted together but did it so without being aware of cooperating. Sharemind brings the "controllers" together but shields their identities from one another; hence, there is no common platform for them to decide and jointly determine the data processing. Thus, according to common legal thinking and notions of "joint partnership," etc., the crucial element for being jointly responsible for a data processing is clearly missing. However, we have to note that the legal discussion has only recently begun, concerning the interpretation and the necessary elements of "joint controllership"– in particular, if this notion has to be interpreted in the same way as traditional partnerships (like in corporate law). Currently, some sort of acting together is still required. Hence, in the case of Sharemind, the users will not be considered as "joint controllers," but rather as the controller for each section of data processing (which is difficult to handle regarding the obligations of data controllers).

However, we have to note that the ECJ ruling on the Google Spain case that was cited above had a very broad notion of joint controllership and failed to deliberate the elements more in-depth. Thus, we are currently confronted with legal uncertainty - in a "worst-case scenario," we should take into account the fact that all users (clients) of Sharemind will, eventually, have to be considered as joint controllers, so that they are each responsible for actions to be taken, in light of the DPD.

[44] See also the legal assessment of Sharemind of the Estonian Data Protection Agency, available at: `http://adr.rik.ee/aki/dokument/4797050`.

[45] See also *Gola/Klug/Körffer*, in: *Gola/Schomerus*, Bundesdatenschutzgesetz, par. 3 recital 30.

3.2.1.2.2 Applicability of Data Protection Law

Multiparty computation is advantageous due to the fact that simply random fragments of personal data are used. The original data can only be restored (and thus turns into personal data) if all fragments are put together. Hence, it is crucial to determine whether the DPD is applicable to the computation over data fragments. The division of data cannot be treated as a traditional form of encryption. Thus, the controversy regarding the sufficient level of encryption is not relevant either.

The qualification of split data is still new to data protection law, yet, well-known in intellectual property law. By splitting works protected by copyright into ‘chunks’, as in peer-to-peer sharing, people try to circumvent the protection of the work provided by copyright law. Although there are differences between copyright law and data protection law, [46] one important parallel can be drawn. For example, if a copyright protected work is split in many parts, and those parts can be perceived, the copyright protection still affects the single parts. The single page of a book, for example, is protected just like the whole book. If the chunks of a copyright-protected work cannot be used for perception of the work (as is the case if archive files like .zip or .rar files are shared via peer-to-peer sharing for example) without having all other parts of the work, some authors argue that the copyright protection does not apply for a single part. [47]

This idea can be used to evaluate if one part of a secret-shared file is still personal data. Without the other two parts, this file cannot be read in any way. One fragment itself does not contain information regarding a person and should not be seen as personal data. For someone looking for information about a certain person, this fragment would be useless. Only if all fragments of the data were gathered and put together the directive would be applicable. Theoretically, all server providers may collude and re-engineer the personal data. However, this is highly unlikely since the providers of the server, themselves, have a high interest in ensuring safety and confidentiality of Sharemind and should be legally bound by contract. Once again, from the stance of the relative approach, the unreasonable chance of collusion leads to ruling out the applicability of the Data Protection Directive. Concluding, we have to point out that the perspective from an absolute approach would differ in taking these chances into account, thus applying the Data Protection Directive.

So, we have to analyze the legal situation in the sense of a “worst-case scenario” if the absolute approach would prevail.

3.2.1.2.3 Compliance with Data Protection Law Now and in the Future

Compliance with the DPD As outlined already for HANA, it is highly recommended to obtain the data subject’s (explicit) consent; if not, there has to be a specific legal permission for processing the data such as fulfillment of contractual obligations.

The consent given must be informed and given with free will on the basis of sufficient information; the same criteria as above apply. If the servers are in a third country, an adequate level of protection has to be guaranteed by the controller using safeguards, as described above. If the servers are within the jurisdiction of the Directive, a processing on behalf of the controller, in the sense of Article 17 of the Directive, could be ensured by a legal framework between the data donors, as controllers, and the Sharemind-provider, as the processor, (“order processing”).

Especially for Sharemind, it is important that the output produced by the multiparty computation cannot be easily re-identified, since the user of Sharemind might be an entity completely different than the original data donors. For instance, data donors can be persons which are interviewed in a

[46] Copyright law aims to protect the right holder against unlawful reproduction of his work whereas data protection law protects the data subject’s right to decide what is done with its personal data.

[47] With further references regarding the copyright-law based discussion concerning illegal sharing of chunks see Heckmann/Nordmeyer, CR 2014, 41 (43).

statistical query, such as students etc., a user can be a research (or commercial) organization that would use the data in order to combine with other data which stem from other parties. In such a scenario, the "user" would be qualified as a data controller. Moreover, further provider may also be involved, which raises the issue of joint controllership again; for instance, a provider of statistical methods and software who determines the data processing. Furthermore, the computing parties have to check that secret data is not published or made accessible by mistake, thus creating a joint controllership. The user (data controller) should be bound by contract to refrain from using Sharemind in a way that would reveal information about the persons which the data donors provided in the first place (for example, combining the donor‘s data with other data derived from Sharemind, thus creating new personal data). Considering the principle of Privacy by Design, this issue should be solved when setting up the contractual framework needed for Sharemind by forbidding the re-identification of persons using contractual penalties as an organizational measure as required by the law (see 2.4.4.1). In order to deter the user of Sharemind from re-identification, the use of auxiliary information to re-identify data subjects and the consequences of doing so have to be disproportionate (unreasonable) in comparison to the value of the personal data they would produce.

Compliance with the GDPR The GDPR will only be applicable to the data processing carried out by the participants of Sharemind if the secretly shared data would be considered 'personal data'. Due to the differences between traditional encryption and secret sharing, it is unlikely that a single part of a secretly-shared data enables the identification of a person, as more than a single key is needed to decrypt the data and they are distributed among different entities with strong interests in keeping the data confidential. A collusion of those parties is highly unlikely. Therefore, we do not think that the upcoming Regulation will be applicable to the computation over secretly shared data, even under the assumption that an absolute approach may prevail under the GDPR (see 2.2.3). To fully assess the possible legal risks the GDPR's main issues with regard to Sharemind shall be described in the following.

The role of the controller is not assigned to one of the participants; rather, it may change for every case Sharemind is used. Even two joint controllers are possible. To ensure a lawful processing under the GDPR all processors involved should process the data on behalf of the respective controller (the user, the client). Hence, before Sharemind is used the parties involved should enter into contractual relations ensuring the requirements described under 2.3.2 are met. It is the controllers' responsibility to bind all other participants legally (in the sense of a contract) and to ensure the necessary technical and organizational measures are implemented. Again, a certification of the processing parties as described in 2.3.2.4 is recommended.

If an 'order processing' takes place compliant to Article 22 ss. of the GDPR the controller has either to obtain the consent of affected persons or to benefit from an explicit legal permission. Therefore, like under the DPD (see 3.2.1.2.3) it is highly recommended to obtain the data subjects consent (see 2.4.2.4).

Akin to the DPD, the GDPR addresses, like the DPD, joint controllership. Both controllers are responsible for the use of Sharemind. They will be bound by Article 26 GDPR to enter into an arrangement that clarifies each controllers' duties, e.g. the information of the supervisory authority in case of a data breach (Art. 33 GDPR), or eventually (if necessary) the appointment of an data protection officer. The arrangement has to be made available to the data subjects, so they can know to whom they can turn if they want to exercise their rights according to Article 17 GDPR (see 2.5.3). Sharemind makes it possible for computation to be carried out over data without the computing parties learning it. If it is ensured that the system's output cannot be re-identified without disproportionate efforts, then the goals of Privacy by Design in Article 25 GDPR can be met.

3.2.2 Secure Collaborative Statistics in Credit Rating

3.2.2.1 Functions

3.2.2.1.1 The Basic Concept

Secure Multiparty Computation (MPC) can be used to facilitate complementary decision support in a traditional credit rating. This business case involves small- to medium-sized Danish banks and an accounting firm. They merge their confidential data by using MPC to create a database. An implemented MPC-based LP-solver is used to compute relative performance analysis of the bank's customers directly on the secretly shared data set. Thus, traditional credit rating can be complemented by means of relative performance evaluations. It is difficult for banks to obtain traditional accountancy information on 'peer' farms, since most farmers are not required to publish and disclose such information (unlike many other businesses). However, their accountancy information which is processed by an accounting firm can be used by involving this accountancy firm in a way that shares the information with banks, without giving them direct access to the personal information of the farmers.
The relative performance analysis of the farms is computed using linear programming, which is one of the most basic and most useful optimization tools. It is widely used in operational research and applied micro economics.
Like in the other use cases mentioned here, none of the involved parties is required to disclose their data to others. Once again, a trusted third party may solve the problem, such as credit scoring agencies in Germany. However, such a solution may turn out too expensive or, on a smaller business scale, too complex to reach.
Another solution is - again - provided by Secure Multiparty Computation (MPC) which allows two or more parties to compute any function without leaking any additional information, other than the output of the function. In this scenario, an LP-solver using MPC primitives has been implemented. Instead of a third trusted party, the MPC coordinates the private information according to a comprehensive protocol; the MPC is used like a trusted third party. In contrast to a trusted third party, MPC does not require one entity to learn all inputs. As the parties involved in MPC are interested in keeping their data confidential, the risk of a privacy breach is lowered compared to a coordination by an uninvolved third party. Banks and the accountancy firm thus provide confidential data without the other party learning these data. An MPC-based LP-solver is used to compute over the datasets to produce a relative performance analysis of the bank's clients– in this case, 'peer' farms. This information can be helpful in evaluating the credit rating of a farm, as well as for evaluating the bank's portfolio of farms.
In the basic scenario, two parties (a bank and an accountancy firm) hold specific data that the other party shall not learn. As long as data is just stored, no encryption of the data is needed to prevent the other party from discovering it because solely the party holding the data can access it. Only when the secure computation of the benchmark takes place the data will be secretly shared between the two servers.
Every farm has a 'CVR Number' - an identifier provided by the Danish Central Business register, i.e. the central register containing primary data on all businesses in Denmark. The bank will be able to obtain a performance analysis of its client - a specific farm that has either been a client before or asks the bank for a loan - by providing the CVR Number to the software. Only the bank knows which farm is one of its clients.

3.2.2.1.2 The Systems Output: Secure Complementary Credit Ranking

The reason a bank may be interested in using this system is that reliable credit scoring and evaluation of a farm is needed in order to meet the banking regulations, concerning great amounts of lending.

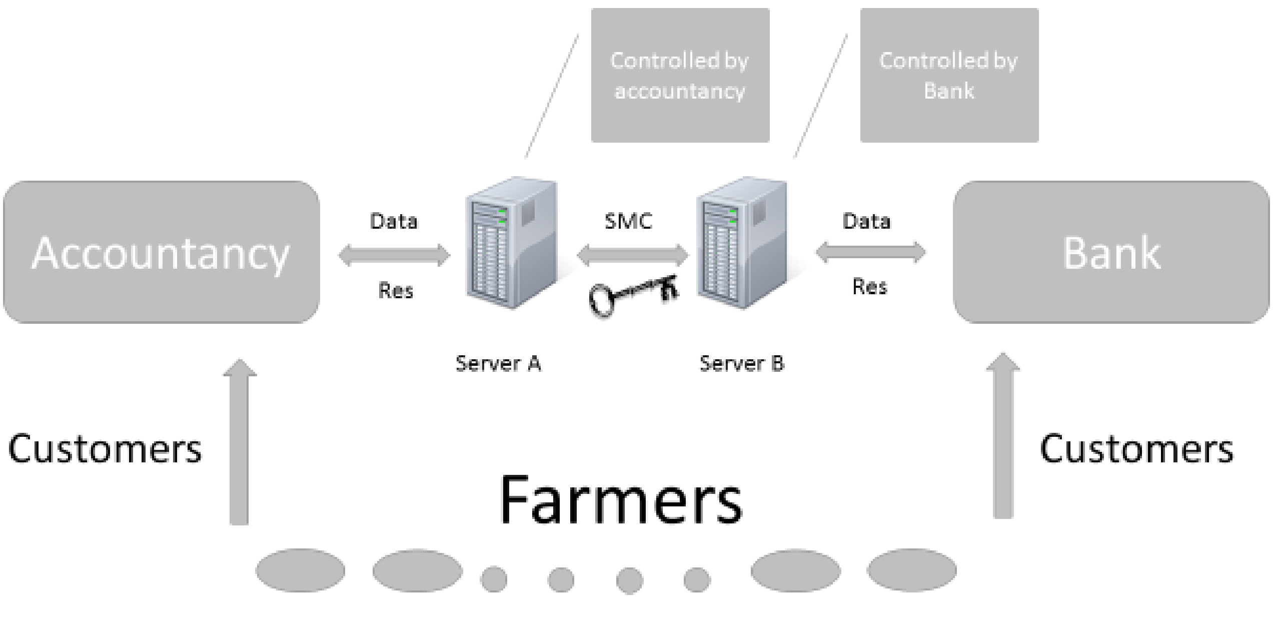


Figure 3.3: The basic functioning of secure collaborative statistics in credit rating using linear programming

Since a bank only has information (typically limited information) about its own customers, smaller banks may lack sufficient data to conduct a proper credit rating analysis of its customers. The software provides a complementary analysis of the firm's relative economic performance, instead of trying to predict the risk of failure in fulfilling its financial commitments to the bank (ie. repaying a potential credit). The bank realizes that this is an efficient scoring method which measures the performance of a farm against its most important peers. The application uses benchmarks (ie. comparing the performance of one unit against that of best practice) with the Data Envelopment Analysis approach. [48] Data Envelopment Analysis can be formulated as an LP-problem. The performance analysis is presented to the bank as a single value without any information that would reveal further information about the farm.

3.2.2.1.3 The Possible Variations of the System

In a straightforward scenario, each involved party is using an Amazon EC2 as a server under the assumption that Amazon gives them full and exclusive control of their respective EC2 instance. However, the bank and the accountants may either use another cloud provider's service or their own IT-resources.
Moreover, the system can be extended to more parties. For instance, instead of having server A run by the accountants house and server B run by a bank, the Danish Bankers Association could run server B, so that every bank that is part of the association, could use server B provided by the association. Since the information is confidential, banks would not want the association to learn this information.

[48] A frontier-evaluation technique that supports best practice comparisons in a multiple-inputs multiple-outputs framework, see *Charnes/Cooper/Rhodes*, European Journal of Operational Research 1978, 429 ff.; *Charnes/Cooper/Rhodes*, European Journal of Operational Research 1978, 339.

Once again, this problem can be solved by secretly sharing the banks data between the two servers. None of the controlling parties of the servers would thus have knowledge of the bank's data. Trust is based on the assumption that the involved parties will not collude – the basic principle of secret sharing. The difference to the basic model described above (see 3.2.2.1.1) refers to the secret sharing of data even when the data is simply stored - in contrast to secret sharing only when data is being computed. The computing would be done by MPC, not by the Banker's Association, or a trusted third party.

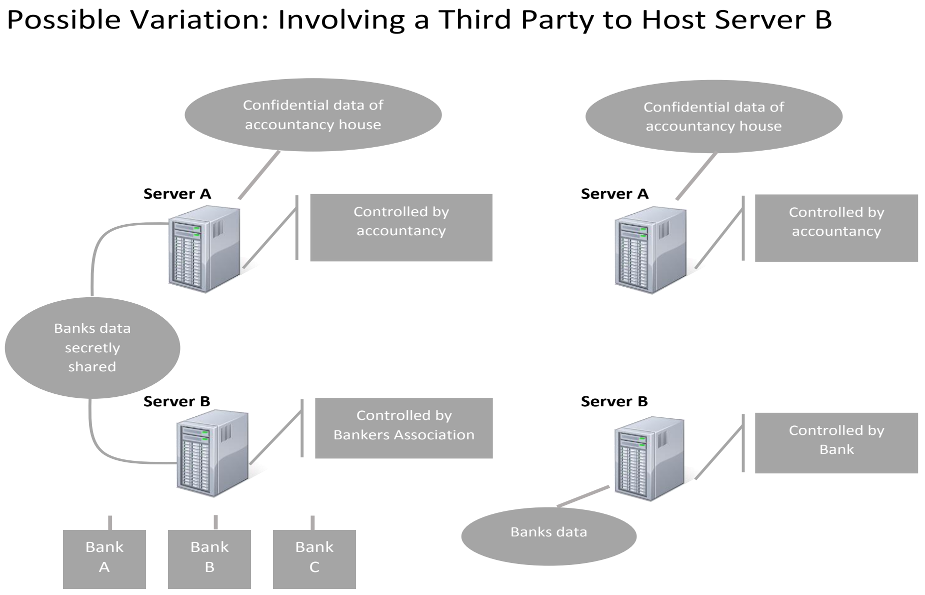

Possible Variation: Involving a Third Party to Host Server B
Confidential data of accountancy house
Server A
Controlled by accountancy
Banks data secretly shared
Server B
Controlled by Bankers Association
Bank A
Bank B
Bank C
Confidential data of accountancy house
Server A
Controlled by accountancy
Server B
Controlled by Bank
Banks data

Figure 3.4: Possible Variation of the basic principle by using a third party to host server B

3.2.2.2 Legal Evaluation and Risk Assessment

3.2.2.2.1 A Legal Classification of the Involved Parties and the Data Processing Activities

The parties involved are the accountancy firm, the bank, the farms, and, eventually, the Danish Bankers Association and Amazon. The bank and the accountancy firm both determine the purposes and means of the processing, as they decide which data will be computed. Therefore, as joint controllers, they are mutually responsible for the processing of the data, Article 2 (d) DPD. If the servers are hosted using Amazon EC2 Instances, then data will be transferred to Amazon and the MPC will be run based on Amazon's Instances. Evidently, Amazon does not determine the purposes of the data processing and, therefore, is not qualified as a controller, but rather as a processor. The same would apply if the Danish Bankers Association would provide the hosting service. If the accountancy firm and/or the banks use their own servers they have to be considered to be controllers and processors at the same time.
In the straightforward scenario, there are four relevant data processing activities: the transfer of the data to the servers, the storage of the data on the two servers, the computation of the data stored on the two servers via MPC, and the production of the performance analysis. If Amazon EC2 Instances are used to run the two servers, a transfer to Amazon as a third party is needed. If the Amazon Instances are hosted on servers outside the EEA/EU, a transfer to a third country is implied. If the two servers are run on physical machines controlled by the bank and the accountancy firm, [49] the transfer of the data to the servers and the storage are not relevant since the entity storing the data will be the entity who is controlling the processing in the first place. This processing would be carried out internally and thus does not face legal challenges.
If the Danish Bankers Association is involved in a transfer of the bank's data to the server run by the association, then the same criteria and principles as for Amazon can be applied.
As explained earlier, the dividing of the data in plaintext cannot be considered as data processing (see 3.2.1.2.1).
Whereas the MPC runs over secret shared data that existed before, the computation of the system's outputs produces new data. If those values are to be considered personal data, the banks using the system would be collecting personal data. This, as well, is data processing that requires an explicit legal permission or the consent of the affected person, here the farmer.

3.2.2.2.2 Applicability of Data Protection Law

The Data Protection Directive only regulates the processing of personal data, which is data that relates to an identified or identifiable natural person. A natural person is a 'normal' human being, and not a company, a corporation or an association. Those are 'legal persons' by law. The processing of data concerning legal persons is not affected by the European Data Protection Directive:

> Recital 24 DPD: "(24) Whereas the legislation concerning the protection of legal persons with regard to the processing data which concerns them is not affected by this Directive"

The data processed in this scenario affects farms accounting and production data, their identities and valuations of their assets. If those farms are organized as companies - in the sense of a legal entity, a legal person (i.e. not run just by a single farmer as a private individual or as a partnership), their data can hardly be considered to be personal data. Nevertheless, the business case deals with small farms and the system described might be useful for the financial performance evaluation of such farms (because there is little to no accountancy data of these farms publicly known). Most of them are privately owned and are not organized as a legal person/entity. However, Denmark opted to include

[49] Provided that the servers of banks and accountancy firms are based in the EU.

legal persons' data in the scope of its data protection law. [50] Moreover, courts in Germany applied the DPD to legal entities run by a single natural person (one-man-company) in the case the data reflects information of the natural person. [51] Hence, there are substantial reasons to assume that natural persons may be affected, specifically, in the business case of farmers and bank loans. Therefore, European data protection law should be taken into account when the system described is applied.

As in the Sharemind case, data will be secretly shared for secure collaborative statistics in credit rating. The secretly shared data fragments are no longer giving away information without the values stored on the other server. Due to the opposing interests of the involved parties (banks and the accountancy firm), it is highly unlikely that a party will learn the other party's information. As showed above, secretly shared information should not be considered to be personal data anymore (see 3.2.1.2.2) - taking also into account the absolute approach.

In contrast, however, to the Sharemind case study not all data are secretly shared all the time. The data is stored on the two servers in plaintext so that all processing concerning this data have to comply with the DPD.

Another disparity with the Sharemind case study refers to the quality of the system's output: whilst with Sharemind, a researcher is not able to re-identify the affected persons and the results are supposed to be anonymized data (see 3.2.1.2.3.1) the case of farmers and banks refers to information being provided for a certain farm: The bank wants to obtain a financial performance scoring in order to use it in addition to the traditional credit rankings. At least for the bank who is aware which farm is involved these results are clearly personal data. Therefore, the data protection law is applicable by way of the collection of personal data. If an absolute approach would be applied to the definition of personal data, the performance analysis - a single value without any identifiers - would have to be considered personal data, even for persons who do not know to which farm this analysis refers.

3.2.2.2.3 Compliance with Existing and Future Data Protection Law

Compliance with the DPD Under the assumption that the data protection law is applicable to the business case both the absolute and the relative approach will lead to the same outcome for some processing activities. Thus, compliance with the DPD has to be ensured.

Using Amazon Instances to host the two servers raises the problem of data protection for the transfer of the data to the servers. The data will be stored in plaintext on those servers and only be secretly shared during the MPC. Hence, even by following the relative approach, this data has to be considered personal data. Since Amazon will be considered to be a third party, the bank and the accountancy firm would need the affected farms' consent or a legal permission for the transfer to Amazon. The consent must be given freely, incorporating all requirements mentioned before (see 2.4.2.3).

In addition, as the system would be used by the banks to assess the credit worthiness of a farm, the transfer of the data can be justified on the grounds of Article 7 lit. (b) or (f) DPD: the processing (here in the form of transferring) can be necessary for the performance of a contract (the loan contract between the farm and the bank) or necessary for the legitimate interests pursued by the bank. As the controller -the bank solely has an interest in lending money to farms financially stable enough to loans it back on-time and with interest.

However, those legal permissions require a balance of interests; the interest of a farm in not having its personal data transferred to a third party may outweigh the interest of the bank in using the system run through Amazon.

The same results for the bank could be produced if the bank and the accountancy firm would run the servers on their own physical machines, without including Amazon. For the accountancy firm,

[50] *Gola/Klug/Körffer*, in: *Gola/Schomerus*, Bundesdatenschutzgesetz, par. 3 recital 11.

[51] *BGH*, NJW 1986, 2505.

which also would be transferring personal data to a server run through Amazon the interest would be even weaker in comparison to the affected farm's interest, since the accountancy firm would not gain a direct advantage (other than making money out of the service or that the bank recognizes the most efficient farms, which would come at the expense of the less efficient farms and therefore might possibly weaken the accountancy firm's interest even further). Thus, it is hard to justify a transfer to Amazon on the grounds of Article 7 lit. (b) or (f).

Moreover, in this version of the scenario, Amazon could hardly be considered a processor on behalf of the controller (see 2.3.1.4), since there would be no contract legally binding Amazon to process data. The processing done with MPC would only take place on Instances run on Amazon servers with Amazon doing nothing more than providing the cloud infrastructure for the system. The controllers would not benefit from the privileged status of a processing carried out on behalf of the controller.

In addition, the transfer to servers outside the EU requires a specific consent by the affected farm or an explicit legal permission concerning the transfer of data to a receiver under the jurisdiction of a third country (see 2.4.3.1). The Amazon Instances may be run on physical machines within the USA, which would allow a transfer on the grounds of the principles for data transfers to third countries (see 2.4.3.1) - however, incorporating all problems already mentioned (2.5.1).

These legal problems can be avoided if the bank and the accountancy firm use physical machines of their own (located in the EU). In this case, there is no transfer to a third party. The only data processing other than the MPC would be the storage on the bank's and the accountancy firms' own servers. The affected farms' interest in not having its data transferred to a third party would not have to be considered in the balance of interests in this scenario. It is more likely that the banks interest in a valuable financial performance analysis before providing a loan to a farm will outweigh the farms interests in that case (note the still weaker interest of the accountancy firm, see above). From a legal perspective the basic scenario using Amazon EC2 Instances clearly involves risks that can be avoided. Still, encryption of the data before transferring it to Amazon may solve the problem if Amazon is to be involved as the DPD will not be applicable to the transfer (according to the relative approach, see 2.2.1).

If the Danish Bankers Association is to be involved (see 3.2.2.1.3), the data transfer to the association has to be evaluated. Since the Association is not allowed to learn the bank's information and would only be used as a means to simplify the system, if more than one bank would want to use it, then the bank's data will be secretly shared between the two servers (one controlled by the Association, one controlled by the accountancy firm). Hence, the two servers have to be treated like the data mining servers in the Sharemind use-case (for the banks data), see 3.2.1.2.3.1. Therefore, neither the storage nor the computation of the bank's data would require a legal permission or consent by the affected farm. It is highly unlikely that the two parties controlling the servers would collude so that one party cloud learn the other party's information in this scenario. If we assume a worst-case scenario where even storing and computation over secretly shared data would fall under the scope of European data protection law, the Association can be considered a processor on behalf of the controller, in this case the banks. The Association would run one of the used mining servers for the storage and for the MPC; however, the banks would still decide over the purposes and means of the processing. An 'order processing' by the Association would be possible if an appropriate contract would be drawn up. The difficulties arising from a cloud-provider functioning as a processor would not occur if the Association (as a 'normal' processor) would process the data. The legal requirements could be met (see 2.3.1.4.3) so that 'order processing' could be assumed. As an organizational measure (see 2.4.4.1) the Association should also agree to an enforceable non-collusion clause in the contract regulating the order processing.

Concerning the system's output the bank will be able to obtain a scoring value that provides information about the financial performance of a certain farm, compared to the performance of other farms.

The bank provides the farms CVR number for the system. The entity who knows to which farm the CVR number belongs to also knows which farm the systems' output value was computed for. Therefore, this value should be considered personal data (if the farms' data is regarded as personal, see 3.2.2.2.2). According to the absolute approach (see 2.2.1) the output value has to be considered to be personal data for everyone. Neither will the output-value be encrypted nor will it be secretly shared, thus differing from the data processing carried out via MPC to produce the output or the secret sharing of the data if the Association is involved. The system's purpose is to produce identifiable data as an output, so safeguards have to be taken in order to ensure compliance with the data protection law. Since the output-value is new data, its production ought to be considered as collection of personal data for the entity retrieving it. For the two entities in charge of producing the output (the accountancy firm and the bank jointly) the production and the provision of the output to the bank has to be considered a transfer of personal data (note that the bank is both one of the joint controllers and the entity receiving the data in the basic scenario). Both the accountancy firm and the bank are jointly responsible for this data processing to be compliant with data protection law; hence, they need, once again, the consent of the farmer or an explicit legal permission, as they do for storing the data on the servers (if it is stored in plaintext). The same result would be reached according to the relative approach.
However, consent may be more easily obtained from farmers asking for a loan without any cloud-specific problems; unlike in the Sharemind case, there is no greater number of affected individuals whose consent would be needed but rather one single farm. Even without consent, the collection of the value might be based on Article 7 lit (b), (f) DPD. Nevertheless, as accountancy firms do have weaker interests in regards to data protection compared to farms (see above), obtaining consent would be the favorable legal option.
A higher risk refers to the potential abuse of the system. Especially in the scenario involving several banks and the Bankers Association a bank may use the CVR number of a farm that is not asking for a loan and who is not one of the banks clients to produce a financial performance evaluation of the farm. The collection of data affecting a farm that is not a client of the bank would be illegal, as there is no legal ground for the collection without any contractual performance etc. The eventual abuse raises a legal risk for the (joint) controller of the system that should be addressed by including technical and organizational safeguards (such as contractual cases for indemnization) before the system is put in use.

Compliance with the GDPR As explained under 3.2.1.2.3.2, the GDPR is not applicable to the processing of secretly shared data. However, the case of Danish banks and farmers implies the storage of data in plaintext, which thus may be used to produce personal data as an output to its user. Therefore, an assessment of the GDPRs impact on this financial-performance-analysis system is needed.
If Amazon EC2 Instances are used to host the two servers, a data transfer to the USA (if the physical machines those Instances are hosted on are seated there) can be legal according to the principles of data transfers to third countries, but with the legal uncertainties described in 2.4.3. If those agreements do not provide legal grounds for the data transfer, Amazon can apply for a certification – the proposed European privacy seal. This seal can enable the banks to provide evidence that they have ensured that the entity (Amazon) in a third country they are transferring data to provides for an adequate level of data protection. [52] Amazon will do no processing in the described scenario (see 3.2.2.2.3.1) and, therefore, no 'order processing' as regulated in Article 22 and the following GDPR will take place.
The situation changes if the Danish Bankers Association is running one of the two servers since the Association will process data via MPC and will secretly share data between the two servers together with the accountancy frim. According to the GDPR, the Association will process the data on behalf of the banks (the controllers). Therefore, as described under 3.2.1.2.3.2 the requirements of Article

[52] Note that this is only the second step, the transfer of data itself has to be lawful, too. See 2.3.4.

22 GDPR have to be respected. Again, a certification of the processor (the Association) in form of the data protection seal is recommended.

In any case, there will be joint controllers, either the accountancy firm and one bank or the accountancy firm and all participating banks, each using the Association as a processor to host their server. Therefore, a contractual framework has to be entered by each partner regulating responsibilities and duties of the involved parties (see 2.3.2.2). [53]

The system will provide an output that has to be considered personal data under the assumption that the affected farms are data subjects. This personal data has not existed before, but it will be newly created data. The financial-performance analysis, therefore, will be a collection of data, according to Article 4 lit (3) GDPR. Hence, according to Article 14 GDPR, the affected farm has to be informed, before the system is put in place. To fulfil Article 14 GDPR's requirements (especially Article 14 Par. 1 lit (f) of the Commission's and of the LIBE-proposal; Article 14 Par. 1a. (c) of the proposal of the Council), it is once again recommended not to use Amazon EC2 Instances to host the two servers.

3.3 Conclusion

To sum up, Part 1 of the deliverable examines the legal challenges of European data protection law regarding cloud computing and encryption, as well as the technologies developed by PRACTICE. Especially the new General Data Protection Regulation, which will come into force on 25 May 2018, raises complex new issues regarding cloud computing. The question whether the DPD is applicable when personal data is encrypted is not solved yet, however, the EJCs decision regarding the personal reference of dynamic IP addresses can be interpreted as a rather relative approach concerning this issue. The GDPRs definition of personal data considers supplementary knowledge of third persons, however, we examined that encrypting personal data using state of the art encryption techniques can in many cases be a way to anonymize personal data with the limitation that a potential possibility of obtaining the key, also by a third party and especially due to decryption, always has to be considered, but only if those means used are reasonably likely to be used. Moreover, the GDPR provides several new obligations for controllers and processors e.g. regarding informed consent of data subjects, transparency or data breach notifications. Applying state-of-the-art encryption technologies can exempt the controller to communicate a personal data breach to the data subject according to the GDPR. A lawful way to process personal data in the cloud both in the DPD as in the GDPR is using order processing on behalf of the controller. Additionally, certifications may be used to demonstrate compliance with the GDPR of processing operations by controllers and processors. Data transfer to third countries is possible by using inter alia instruments such as standard data protection clauses adopted or approved by the EU Commission, data protection certifications or binding corporate rules. After the judgment of the ECJ declaring the Safe Harbor agreement to be invalid, the transfer of personal data to the U.S. can now be based on the EU-US Privacy Shield, the succeeding legal framework. However, the compliance of the EU-US Privacy Shield with the GDPR and the European fundamental rights remains uncertain, as the new framework does not consistently increase the level of data protection for data subjects. Furthermore, the technologies developed by PRACTICE comply ideally with the new principle of Privacy by Design. If Encrypted HANA and Sharemind are used, according to the relative approach the encrypted data would not be qualified as personal data.

[53] As described above: the information of the supervisory authority in case of a data breach (Art. 31 GDPR), if needed the appointment of a data protection officer (2.5.6), also the making available to the data subjects (the farms), in case they want to exercise their rights following Art. 17 GDPR (see 2.5.5).

List of Abbreviations

Abbreviation	German spelling	English spelling
AG	Amtsgericht	District Court
BB	Betriebs Berater	Operation advisor (journal)
BCR	-	Binding Corporate Rules
BDSG	Bundesdatenschutzgesetz	German Federal data protection act
BeckRS	Beck-Rechtsprechung	Beck-jurisdiction
CFR	-	Charter of Fundamental Rights of the European Union
CR	Computer und Recht	Computers and Law (journal)
DPA	-	Data Protection Authority
DPD	-	Data Protection Directive
DuD	Datenschutz und Datensicherheit	Data protection and data security (journal)
ECJ	-	European Court of Justice
ENISA	-	European Union Agency for Network and Information Security
EuZW	Europäische Zeitschrift für Wirtschaftsrecht	European journal of Business Law (journal)
GDPR	-	Proposal for a General Data Protection Regulation
GRUR	Gewerblicher Rechtschutz und Urheberrecht	Intellectual property and copyright (journal)
ITRB	Der IT-Rechts-Berater	IT law adviser
JIPITEC	-	Journal of Intellectual Property, Information Technology and E-Commerce Law
jurisPR-ITR	Juris Praxis Report - IT-Recht	Juris practice report - IT-law (online journal)
JZ	JuristenZeitung	Lawyers' Journal (journal)
K&R	Kommunikation und Recht	Communication and Law
KG	Kammergericht	See OLG
LG	Landgericht	Regional court
LMuR	Lebensmittel und Recht	Foodstuffs and law (journal)
LP	-	Linear Programming
MMR	MultiMedia und Recht	MultiMedia and law (journal)
NIST	-	National Institute of Standards and Technology
NJW	Neue Juristische Wochenschrift	New weekly report on legal issues (journal)
OLG	Oberlandesgericht	Higher regional court (or circuit court)
OVG	Oberverwaltungsgericht	Higher administrative Court (circuit court)
RDV	Recht der Datenverarbeitung	Law of data processing (journal)
SCA	-	Stored Communications Act

Abbreviation	German spelling	English spelling
SMC	-	Secure Multiparty Computation
TMG	Telemediengesetz	Telemedia Act
TTP	-	Trusted Third Party
WP	-	Working Party
VG	Verwaltungsgericht	Administrative Court
ZD	Zeitschrift für Datenschutz	Journal of data protection
ZUM	Zeitschrift für Urheber- und Medienrecht	Journal of Copyright and Media Law

Bibliography

[1] Alich, Stefan; Nolte, Georg: Zur datenschutzrechtlichen Verantwortlichkeit (außereuropäischer) Hostprovider für Drittinhalte. In CR, 741 ff, 2011.

[2] Arbeitskreise Technik und Medien der Konferenz der Datenschutzbeauftragten des Bundes und der Länder sowie der Arbeitsgruppe Internationaler Datenverkehr des Düsseldorfer Kreises: Orientierungshilfe - Cloud Computing, Version 2.0. Available at: https://www.datenschutz-bayern.de/technik/orient/oh_cloud.pdf.

[3] Art. 29-Working Party: Opinion 04/2012 on Cookie Consent Exemption, WP 194, 07/06/2012. Available at: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2012/wp194_en.pdf

[4] Art. 29-Working Party: Opinion 05/2012 on Cloud Computing, WP 196, 01/07/2012. Available at: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2012/wp196_en.pdf

[5] Art. 29-Working Party: Opinion 15/2011 on the definition of consent, WP 187, 13/07/2011. Available at: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2011/wp187_en.pdf

[6] Art. 29-Working Party: Opinion 03/2009 on the Draft Commission Decision on standard contractual clauses for the transfer of personal data to processors established in third countries, under Directive 95/46/EC, WP 161, 05/03/2009. Available at: http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2009/wp161_en.pdf

[7] Art. 29-Working Party: Opinion 04/2007 on the concept of personal data, WP 136, 20/06/2007. Available at: http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp136_en.pdf

[8] Art. 29-Working Part: Opinion 08/2010 on applicable law, WP 179, 16/12/2010. Available at: http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2010/wp179_en.pdf

[9] Art. 29-Working Party: Opinion 01/2010 on the concepts of “controller” and “processor”, WP 169, 16/02/2010. Available at: http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2010/wp169_en.pdf

[10] Art. 29-Working Party:Opinion 05/2014 on Anonymisation Techniques, WP 216, 10/04/2014. Available at: http://www.cnpd.public.lu/de/publications/groupe-art29/wp216_en.pdf

[11] Art. 29-Working Party: Update of Opinion 8/2010 on applicable law in light of the CJEU judgement in Google Spain, WP 179 update, 13/12/2015. Available at: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2015/wp179_en_update.pdf.

[12] Art. 29-Working Party: Opinion 01/2016 on the EU - U.S. Privacy Shield draft adequacy decision, WP 238, 13/04/2016. Available at: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2016/wp238_en.pdf.

[13] Art. 29-Working Party: Working Document: Transfers of personal data to third countries: Applying Article 26 (2) of the EU Data Protection Directive to Binding Corporate Rules for International Data Transfers, WP 74, 03/06/2003. Available at: http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2003/wp74_en.pdf

[14] Art. 29-Working Party: Working Document: Setting up a framework for the structure of Bindin Corporate Rules, WP 154, 24/06/2008. Available at http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2008/wp154_en.pdf.

[15] Bär, Wolfgang: Anmerkung zu: BGH: Speicherung von IP-Adressen durch die Bundesrepublik. In MMR, 134 ff., 2015.

[16] Berg, Kay Uwe: EU-Datenschutzgrundverordnung - Das Aus für Auskunfteien und Inkassounternehmen?. In PinG, 69 ff, 2013.

[17] Bergauer, Christian: Indirekt personenbezogene Daten - datenschutzrechtliche Kuriosa. In Jahrbuch Datenschutzrecht, 55 ff, 2011.

[18] Bergemann, Benjamin: EU-Datenschutzverordnung darf nicht Merkles NAS-Feigenblatt werden - Netzpolitik.org. 17/08/2013. Available at: https://netzpolitik.org/2013/eu-datenschutzverordnung-darf-nicht-merkels-nsa-feigenblatt-werden/

[19] Bergt, Matthias: Anmerkung zur Entscheidung des BGH (Beschluss vom 28.10.2014 - VI ZR 135/13, ZD 2015, 80) zur Speicherung von IP-Adressen durch die Bundesrepublik. In ZD, 83 ff., 2015.

[20] Bergt, Matthias: Das Ende der Rechtssicherheit im Datenschutzrecht. CRonline-Blog of 19/10/2016, available at: http://www.cr-online.de/blog/2016/10/19/das-ende-der-rechtssicherheit-im-datenschutzrecht/.

[21] Bergt, Matthias: Die Bestimmbarkeit als Grundproblem des Datenschutzrechts. In ZD, 365 ff., 2015.

[22] Bergt, Matthias: IP-Adressen: EU-Kommission gibt BGH Nachhilfe in Sachen Grundrechte. CRonline-Blog of 13/09/2015, available at: http://www.cr-online.de/blog/2015/09/13/ip-adressen-eu-kommission-gibt-bgh-nachhilfe-in-sachen-grundrechte/

[23] Bergt, Matthias: Der Einfluss des Safe-Harbour-Urteils auf den Entwurf der Datenschutz-Grundverordnung. Available at: http://www.cr-online.de/blog/2016/01/04/der-einfluss-des-safe-harbor-urteils-auf-den-entwurf-der-datenschutz-grundverordnung/

[24] Bitkom: Leitfaden Cloud Computing, 2009. Available at: http://www.bitkom.org/files/documents/BITKOM-Leitfaden-CloudComputing_Web.pdf

[25] Boerding, Andreas: Ein neues Datenschutzschild für Europa - Warum auch das überarbeitete Privacy Shield den Vorgaben des Safe Harbor-Urteils des EuGH nicht gerecht werden kann. In CR 2016, 431 ff.

[26] Bogdanov, Dan: Sharemind: programmable secure computations with practical applications, PhD thesis, University of Tartu, 2013. Available at: http://dspace.utlib.ee/dspace/bitstream/handle/10062/29041/bogdanov_dan_2.pdf?sequence=5

[27] Bogdanov, Dan; Kamm, Liiana; Laur, Sven; Pruulmann-Vengerfedt, Pille: Secure multi-party data analysis: end user validation and practical experiments, 2013. Available at: http://eprint.iacr.org/2013/826.pdf

[28] Borges, Georg; Meents, Jan Geert (ed.): Cloud Computing - Rechtshandbuch. Munich, 2016.

[29] Brennscheid, Kristin: Cloud Computing und Datenschutz, Diss. Bochum, Baden-Baden, 2013.

[30] Brink, Stefan; Eckhardt, Jens: Wann ist ein Datum ein personenbezogenes Datum? Anwendungsbereich des Datenschutzrechts. In ZD, 205 ff., 2015.

[31] Brisch, Klaus: Pieper, Fritz: Das Kriterium der "Bestimmbarkeit" bei Big Data-Analyseverfahren: Anonymisierung, Vernunft und rechtliche Absicherung bei Datenübermittlungen. In CR 2015, 724 ff.

[32] Buchner, Benedikt: Grundsätze und Rechtmäigkeit der Datenverarbeitung unter der DS-GVO. In DuD 2016, 155 ff.

[33] v. d. Bussche, Axel; Voigt, Paul: Konzerndatenschutz - Rechtshandbuch. Munich, 2014.

[34] Charnes, A.; Cooper, W.W; Rhodes, E: Measuring the efficiency of decision making units. In European Journal of Operational Research, 429-444, 1978.

[35] Charnes, A.; Cooper, W.W; Rhodes, E.: Short communication: measuring the efficiency of decision making units. In European Journal of Operational Research, 339, 1978.

[36] Cybernetica: sharemind : Your secure service plattform for data collection and analysis. Available at: https://sharemindSharemind.cyber.ee/files/images/SharemindSharemind%20secure%20service%20platform%202012.pdf

[37] Däubler, Wolfgang; Klebe, Thomas; Wedde, Peter; Weichert, Thilo (ed.), Bundesdatenschutzgesetz - Kompaktkommentar. 4th Edition, Frankfurt/Main, 2014.

[38] Dammann, Ulrich; Simitis, Spiros: EG-Datenschutzrichtlinie - Kommentar. 1st Edition, Baden-Baden, 1997.

[39] Dammann, Ulrich: Erfolge und Defizite der EU-Datenschutzgrundverordnung: Erwarteter Fortschritt, Schwächen und überraschende Innovationen. In ZD 2016, 307 ff.

[40] Decker, Florian: Die neue europäische Datenschutzgrundverordnung - welche änderungen sind für deutsche Unternehmen zu erwarten?, 2013. Available at: http://blog-it-recht.de/2013/12/02/die-neue-europaeische-datenschutzgrundverordnung-welche-aenderungen-sind-fuer-deutsche-unternehmen-zu-erwarten/

[41] Drews, Stefan; Moneal, Manfred: Grenzenlose Auftragsdatenverarbeitung. In PinG, 143 ff., 2014.

[42] Düsseldorfer Kreis: Beschluss der Aufsichtsbehörden für den Datenschutz im nicht-öffentlichen Bereich vom 13./14. September 2016: Fortgeltung bisher erteilter Einwilligungen unter der Datenschutz-Grundverordnung. Available at: https://www.lda.bayern.de/media/dk_einwilligung.pdf

[43] Eckhardt, Jens: Kommentar zu: LG Berlin, Urteil vom 06.09.2007 - 23 S 3/07. In K&R, 601 ff., 2007.

[44] Eckhardt, Jens: IP-Adresse als personenbezogenes Datum - neues Öl ins Feuer. In CR, 339 ff., 2011.

[45] Eckhardt, Jens; Kramer, Rudi; Mester, Brita Alexandra: Auswirkungen der geplanten EU-DS-GVO auf den deutschen Datenschutz. In DUD, 623 ff., 2013.

[46] Eckhardt, Jens: Cloud Computing - Orientierungshilfe 2.0 des Düsseldorfer Kreises. In DuD, 176 ff., 2015.

[47] Ehmann, Eugen; Helfrich, Marcus: EG-Datenschutzrichtlinie - Kurzkommentar. 1st. Edition, Cologne, 1999.

[48] ENISA: Privacy and Data Protection by Design - from policy to engineering, 2014. Available at: https://www.enisa.europa.eu/publications/privacy-and-data-protection-by-design

[49] Esayas, Samson Yoseph: The role of anonymisation and pseudonymisation under the EU data privacy rules: beyond the 'all or nothing' approach. In European Journal of Law and Technology (2015) Vol 6, No 2, 1 ff.

[50] Faust, Sebastian; Spittka, Jan; Wybitul, Tim: Milliardenbußgelder nach der DS-GVO? Ein Überblick über die neuen Sanktionen bei Verstößen gegen den Datenschutz. In ZD 2016, 120 ff.

[51] Fazlioglu, Muge: Forget me not: the clash of the right to be forgotten and freedom of expression on the Internet. In International Data Privacy Law, p. 149 ff., 2013. Available at: http://idpl.oxfordjournals.org/content/3/3/149.full.pdf+html

[52] Forgó, Nikolaus: My health datayour research: some preliminary thoughts on different values in the General Data Protection Regulation. In International Data Privacy Law (2015) Vol. 5, No. 1, 54 ff.

[53] Frauenhofer Institut für Offene Kommunikationssysteme: ISPRAT-Studie, Cloud-Computing für die öffentliche Verwaltung, 11/2010. Available at: http://www.cloud.fraunhofer.de/content/dam/allianzcloud/de/documents/ISPRAT_cloud_studievorabversion20101129tcm421-76759.pdf

[54] Funke, Michael; Wittmann, Jörn: Cloud Computing - ein klassischer Fall der Auftragsdatenverarbeitung?. In ZD, 221 ff., 2013.

[55] Gerlach, Carsten: Personenbezug von IP-Adressen. In CR, 478 ff., 2013.

[56] German Federal Office for Information Security Technology: BSI-Standard 100-1

[57] German Federal Office for Information Security Technology: BSI-Standard 100-1 Information Security Management Systems (ISMS). Available at: `https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/Publications/BSIStandards/standard_100-1_e_pdf.pdf?__blob=publicationFile`

[58] German Federal Office for Information Security Technology: Safety Recommendation for Cloud Computing Providers. Available at: `https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Mindestanforderungen/Eckpunktepapier-Sicherheitsempfehlungen-CloudComputing-Anbieter.pdf?__blob=publicationFile`

[59] Giedke, Anna: Cloud Computing: Eine wirtschaftsrechtliche Analyse mit besonderer Berücksichtigung des Urheberrechts. Diss., Munich, 2013.

[60] Gierschmann, Sibylle; Was bringt deutschen Unternehmen die DS-GVO? Mehr Pflichten, aber die Rechtsunsicherheit bleibt. In ZD 2016, 51 ff.

[61] Gola, Peter; Lepperhoff, Niels: Reichweite des Haushalts- und Familienprivilegs bei der Datenverarbeitung - Aufnahme und Umfang der Ausnahmeregelung in der DS-GVO. In ZD 2016, 9 ff.

[62] Gola, Peter; Schomerus, Rudolf (ed.): BDSG Bundesdatenschutzgesetz - Kommentar. 11th Edition, Munich, 2012.

[63] Hansen, Marit: Datenschutz-Folgenabschätzung - gerüstet für Datenschutzvorsorge? In DuD 2016, 587 ff.

[64] Härting, Niko: Datenschutzreform in Europa: Einigung im EU-Parlament : Kritische Anmerkungen. In CR, 715 ff., 2013.

[65] Härting, Niko: Internetrecht. 5th Edition, Cologne, 2014.

[66] Härting, Niko: Starke Behörden, schwaches Recht - der neue EU-Datenschutzentwurf. In BB, 459 ff., 2012.

[67] Härting, Niko: Schutz von IP-Adressen. In ITRB, 35 ff., 2009.

[68] Härting, Niko: Datenschutz-Grundverordnung - Das neue Datenschutzrecht in der betrieblichen Praxis. Cologne, 2016.

[69] Härting, Niko: Datenschutz-Grundverordnung Anwendungsbereich, Verbotsprinzip, Einwilligung. In ITRB 2016, 36 ff.

[70] Härting, Niko: Auftragsverarbeitung nach der DSGVO. In ITRB 2016, 137 ff.

[71] Heckmann, Dirk (ed.): Juris PraxisKommentar Internetrecht. 4th Edition, Saarbrücken, 2014.

[72] Heckmann, Jörn; Nordmeyer, Arne: Pars pro toto: Verletzung des Urheberrechtsgesetzes durch das öffentliche Zugänglichmachen von Dateifragmenten (“Chunks”) in Peer-to-Peer-Tauschbörsen. In CR, 41-45, 2014.

[73] Heidrich, Joerg; Wegener, Christoph: Sichere Datenwolken Cloud Computing und Datenschutz. In MMR, 803, 2010.

[74] Heinemeyer, Dennis: Verfahrensstand-Anzeiger. In Verordnung des Europäischen Parlaments und des Rates zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten und zum freien Datenverkehr (Datenschutz-Grundverordnung). Available at: `http://www.computerundrecht.de/26378.htm`

[75] Hennrich, Thorsten: Compliance in Clouds. In CR, 546 ff, 2011.

[76] Hennrich, Thorsten: Cloud Computing - Herausforderungen an den Rechtsrahmen für Datenschutz. Diss., Berlin, 2016.

[77] Hilber, Marc: Handbuch Cloud Computing. Cologne, 2014.

[78] Hon, W Kuan; Millard, Christopher; Walden, Ian: The Problem of 'Personal Data' in Cloud Computing - What Information is Regulated?, The Cloud of Unknowing, Part. 1. 10/03/2011. Available at: `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1783577`

[79] Hon, W Kuan; Hörnle, Julia; Millard, Christopher: Data Protection Jurisdiction and Cloud Computing - When are cloud Users and Providers Subject to EU Data Protection Law?, The Cloud of Unknowing, Part 3. 09/02/2012.

[80] Hon, W Kuan; Millard, Christopher; Walden, Ian: Who is Responsible of 'Personal Data" in Cloud Computing?, The Cloud of Unknowing, Part 2. 21/03/2011. Available at: `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1794130`

[81] Hon, W Kuan; Millard, Christopher: Data Export in Cloud Computing; How Can Personal Data Be Transferred Outside the EEA?, The Cloud of Unknowing, Part 4. 04/04/2012. Available at: `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2034286`

[82] Hon, W Kuan; Kosta, Eleni; Millard, Christopher; Stefanatou, Dimitra: Cloud Accountability: The Likely Impact of the Proposed EU Data Protection Regulation. In Tilburg Law School Legal Studies Research Paper Series No. 07/2014. Available at: `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2405971`

[83] Hornung, Gerrit; Sädtler, Stephan: Europas Wolken - Die Auswirkungen des Entwurfs für eine Datenschutz-Grundverordnung auf das Cloud Computing. In CR, 638 ff., 2012.

[84] Hullen, Nils: Anonymisierung und Pseudonymisierung in der Datenschutzgrundverordnung. PinG 2015, 210 ff.

[85] Jandt, Silke; Roßnagel, Alexander: Datenschutz in Social Networks - Kollektive Verantwortlichkeit für die Datenverarbeitung. In ZD, 160 ff., 2011.

[86] Jaspers, Andreas: Die EU-Datenschutz-Grundverordnung : Auswirkungen der EU-Datenschutz-Grundverordnung auf die Datenschutzorganisation des Unternehmens. In DuD, 571 ff., 2012.

[87] Jotzo, Florian: Gilt deutsches Datenschutzrecht auch für Google, Facebook & Co. bei grenzüberschreitendem Datenverkehr?. In MMR, 232 ff., 2009.

[88] Jotzo, Florian: Der Schutz personenbezogener Daten in der Cloud. Diss., Kiel 2013.

[89] Jülich, Tim; Röttgen, Charlotte; v. Schönfeld, Max: Das Recht auf Datenübertragbarkeit - Ein datenschutzrechtliches Novum. In ZD 2016, 358 ff.

[90] Kamm, Liina; Willemson, Jan: Secure Floating-Point Arithmetic and Private Satellite Collision Analysis. Available at: `http://eprint.iacr.org/2013/850.pdf`

[91] Karg, Moritz: Anonymität, Pseudonyme und Personenbezug revisited? In DuD, 520 ff., 2015.

[92] Keppeler, Lutz M.: "Objektive Theorie" des Personenbezugs und "berechtigtes Interesse" als Untergang der Rechtssicherheit? Eine Analyse der Schlussanträge des Generalanwalts in der Rechtssache C-582/14 (Speicherung dynamischer IP-Adressen). In CR 2016, 360 ff.

[93] Kilian, Wolfgang; Heussen, Benno (ed.): Computerrechts-Handbuch: Computertechnologie in der Rechts- und Wirtschaftspraxis. Supplement 32, Munich, 2013.

[94] Kindt, Els J.: Why research may no longer be the same: about the territorial scope of the Proposed Data Protection Regulation. In CiTiP Working Paper Series 26/2016. Available at: `http://papers.ssrn.com/sol3/JELJOUR_Results.cfm?form_name=journalbrowse&journal_id=1781425`

[95] Klar, Manuel: Räumliche Anwendbarkeit des (europäischen) Datenschutzrechts. In ZD, 109 ff., 2013.

[96] Klinger, Markus: Vorschlag zur EU-Datenschutz-Grundverordnung i.d.F. des EU-Parlaments - Auswirkungen auf datenverarbeitende Unternehmen im Überblick. In jurisPR-ITR 6/2014 Anm. 2.

[97] Kokott, Juliane; Sobotta, Christoph: The distinction between privacy and data protection in the jurisprudence of the CJEU and the ECtHR. In International Data Privacy Law, 222 ff., 2013. Available at: `http://idpl.oxfordjournals.org/content/3/4/222.full.pdf+html`

[98] Koós, Clemens; Englisch, Bastian: Auftragsdatenverarbeitung? - Gegenüberstellung der aktuellen Rechtslage und der DS-GVO in der Fassung des LIBE-Entwurfs. In ZD, 276 ff., 2014.

[99] Krebs, David: "Privacy by Design": Nice-to-have or a Necessary Principle of Data Protection Law? In JIPITEC 2013, 2 ff. Available at: `http://www.jipitec.eu/issues/jipitec-4-1-2013/jipitec4krebs/jipitec-4-1-2013-2-krebs.pdf`

[100] Krempl, Stefan: EU-Datenschützer fordert Einbau von Datenschutz in die Technik. Available at: `http://www.heise.de/newsticker/meldung/EU-Datenschuetzer-fordert-Einbau-von-Datenschutz-in-die-Technik-960735.html`

[101] Kroschwald, Steffen: Verschlüsseltes Cloud Computing : Auswirkung der Kryptografie auf den Personenbezug in der Cloud. In ZD, 75 ff., 2014.

[102] Kühling, Jürgen: Auf dem Weg zum vollharmonisierten Datenschutz?!. In EuZW, 281 ff., 2012.

[103] Kühling, Jürgen; Klar, Manuel: Unsicherheitsfaktor Datenschutzrecht - Das Beispiel des Personenbezugs und der Anonymität. In NJW, 3611 ff., 2013.

[104] Kuner, Christopher: Legal Aspects of Encryption in the Internet. International Business Lawyer 1996, 186 ff.

[105] Kuner, Christopher: European Data Protection Law. 2nd Edition, New York, 2007.

[106] Lagos, Yianni: Taking the Personal Out of Data: Making Sense of De-Identification. In Indiana Law Review 2014 - 2015, 187 ff. Available at: https://mckinneylaw.iu.edu/ilr/pdf/vol48p187.pdf

[107] Lang, Markus: Reform des EU-Datenschutzrechts. In K&R, 145 ff., 2012.

[108] Leonard, Peter: Customer data analytics: privacy settings for 'Big Data' business. In International Data Privacy Law, Vol. 4, No. 1, 53 ff., 2014. Available at: http://idpl.oxfordjournals.org/content/4/1/53.full.pdf+html?sid=a11bd260-1434-4d5e-8e3b-c15b73574748

[109] Leutheusser-Schnarrenberger, Sabine: Zur Reform des europäischen Datenschutzrechts. In MMR, 709 f., 2012.

[110] Lundevall-Unger, Patrick; Tranvik, Tommy: IP Addresses - Just a Number? In International Journal of Law and Information Technology 2010, 53 ff.

[111] Maisch, Michael Marc: Nutzertracking im Internet. In ITRB, 13 ff., 2011.

[112] Mantelero, Alessandro: The EU Proposal for a General Data Protection Regulation and the roots of the right to be forgotten. In Computer Law & Security Review 2013, 229 ff.

[113] Marnau, Ninja; Schlehahn, Eva: Cloud Computing: Legal Analysis. In TClouds (D1.2.2). Available at: http://www.tclouds-project.eu/downloads/deliverables/TC-D1.2.2_Cloud_Computing-Legal_Analysis_M12.pdf

[114] Marnau, Ninja: Anonymisierung, Pseudonymisierung und Transparenz für Big Data - Technische Herausforderungen und Regelungen in der Datenschutz-Grundverordnung. In DuD 2016, 428 ff

[115] Marschall, Kevin: Datenpannen - "neue" Meldepflicht nach der europäischen DS-GVO? Rechtliche Änderungen durch Art. 31 und Art. 32 DS-GVO. In DuD, 183 ff, 2015.

[116] Mayer-Schönberger, Viktor; Padova, Yann: Regime Change? Enabling Big Data through Europe's New Data Protection Regulation. In: The Columbia Science and Technology Law Review 2016, 315 ff. Available at: http://stlr.org/download/volumes/volume17/SchonbergerPadova.pdf

[117] Meyerdierks, Per: Sind IP-Adressen personenbezogene Daten?. In MMR, 8 ff., 2009.

[118] Millard, Christopher: Cloud Computing, Oxford, 2013.

[119] Nägele, Thomas; Jacobs, Sven: Rechtsfragen des Cloud Computing. In ZUM, 281 ff., 2010.

[120] NASA: Satellite Collision Leaves Significant Debris Clouds. In Orbital Debris Quarterly News, Volume 13, Issue 2, 1-2, 2009. Available at: http://orbitaldebris.jsc.nasa.gov/newsletter/pdfs/ODQNv13i2.pdf

[121] Nebel, Maxi, Richter, Philipp. Datenschutz bei Internetdiensten nach der DS-GVO - Vergleich der deutschen Rechtslage mit dem Kommissionsentwurf. In ZD, 407 ff., 2012.

[122] Niemann, Fabian; Paul, Jörg-Alexander: Bewölkt oder wolkenlos - rechtliche Herausforderungen des Cloud Computings. In K&R, 444 ff., 2009.

[123] Niemann, Fabian; Ammann, Jörg-Alexander: Praxishandbuch Rechtsfragen des Cloud Computing. Berlin, Bosten, 2014.

[124] Nink, Judith; Pohle, Jan: Die Bestimmbarkeit des Personenbezugs - Von der IP-Adresse zum Anwendungsbereich der Datenschutzgesetze. In MMR, 563 ff., 2015.

[125] Nord, Jantina; Manzel, Manzel: Datenschutzerklärungen- - misslungene Erlaubnisklauseln zur Datennutzung : -Happy-Digits- und die bedenklichen Folgen im E-Commerce. In NJW, 3756, 2010.

[126] Pahlen-Brandt, Ingrid: Datenschutz braucht scharfe Instrumente: Beitrag zur Diskussion um -personenbezogene Daten. In DuD, 34 ff., 2008.

[127] Peifer, Karl-Nikolaus: Verhaltensorientierte Nutzeransprache - Tod durch Datenschutz oder Moderation durch das Recht?. In K&R, 543 ff., 2011.

[128] Petri, Thomas, Auftragsdatenverarbeitung - heute und morgen. Reformüberlegungen zur Neuordnung des Europäischen Datenschutzrechts. In ZD, 305 ff., 2015.

[129] Piltz, Carlo: Datenschutzreform: aktueller Stand der Verhandlungen im Rat. 20/01/2014. Available at: http://www.delegedata.de/2014/01/datenschutzreform-aktueller-stand-der-verhandlungen-im-rat/

[130] Piltz, Carlo: Datenschutzrecht und Webseiten: Welches Recht ist anwendbar und welche Aufsichtsbehörde ist zuständig? Zugleich Anmerkung zum Vorabentscheidungsersuchen C-230/14 - Weltimmo. In K&R 2015, 559 ff.

[131] Plath, Kai-Uwe: Datenherausgabepflicht für Cloud-Anbieter nach US-Recht vs. EU-Datenschutzrecht. 13/05/2014. Available at: http://www.cr-online.de/blog/2014/05/13/datenherausgabepflicht-fuer-cloud-anbieter-nach-us-recht-vs-eu-datenschutzrecht/

[132] Plath, Kai-Uwe (ed.): Kommentar zum BDSG und zur DSGVO sowie den Datenschutzbestimmungen des TMG und TKG. 2nd Edition, Cologne, 2016

[133] Pohle, Jan; Ammann, Thorsten: Software as a Service - auch rechtlich eine Evolution?. In K&R, 625 ff., 2009.

[134] Pollirer, Hans-Jürgen; Weiss, Ernst M.; Knyrim, Rainer: Datenschutzgesetz 2000 (DSG 2000) samt ausführlichen Erläuterungen. 2nd Edition, Vienna, 2014.

[135] Pollmann, Maren; Kipker, Dennis-Kenji: Informierte Einwilligung in der Online-Welt. In DuD 2016, 378 ff.

[136] Polonetsky, Jules; Tene, Omer; Finch, Kelsey: Shades of Gray: Seeing the Full Spectrum of Practical Data De-Identification. Santa Clara Law Review (Forthcoming) 2016, 593 ff. Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2757709

[137] Popa, Raluca Ada: Research Statement. Available at: http://www.mit.edu/~ralucap/researchstatement.pdf

[138] Popa, Raluca Ada; Zeldovich, Nickolai; Balakrishnan, Hari: CryptDB: A Practical Encrypted Relational DBMS. In Technical Report MIT-CSAIL-TR-2011-005, Computer Science and Artificial Intelligence Laboratory, Cambridge, MA, 01/2011. http://people.csail.mit.edu/nickolai/papers/raluca-cryptdb-tr.pdf

[139] Popa, Raluca Ada; Redfield, Cahterine M. S.; Zeldovich, Nickolai; Balakrishnan, Hari: CryptDB: Protecting Confidentiality with Encrypted Query Processing. In Proceedings of the 23rd ACM Symposium on Operating Systems Principles (SOSP), Cascais, Portugal, 10/2011. Available at: http://people.csail.mit.edu/nickolai/papers/raluca-cryptdb.pdf

[140] Rammos, Thanos: Datenschutzrechtliche Aspekte verschiedener Arten ”verhaltensbezogener” Onlinewerbung. In K&R, 692 ff., 2011.

[141] Rath, Michael; Rothe, Britta: Cloud Computing: Ein datenschutzrechtliches Update. In K&R, 623 ff., 2013.

[142] Reding, Viviane: The European data protection framework for the twenty-first century. In International Data Privacy Law 2012, 119 ff.

[143] Roßnagel, Alexander (ed.): Beck‘scher Kommentar zum Recht der Telemediendienste. Munich, 2013.

[144] Roßnagel, Alexander (ed.): Handbuch Datenschutzrecht: Die neuen Grundlagen für Wirtschaft und Verwaltung. Munich, 2003.

[145] Roßnagel, Alexander; Scholz, Philip: Datenschutz durch Anonymität und Pseudonymität : Rechtsfolgen der Verwendung anonymer und pseudonymer Daten. In MMR, 721 ff., 2000.

[146] Roßnagel, Alexander; Richter,Philipp; Nebel, Maxi: Besserer Internetdatenschutz für Europa - Vorschläge zur Spezifizierung der DS-GVO. In ZD, 103 ff., 2013.

[147] Savin, Andrej: EU Internet Law. Cheltenham, 2013

[148] Schaar, Peter: Privacy By Design. Available at: http://www.bfdi.bund.de/SharedDocs/Publikationen/EN/0610EUPrivacyByDesign.pdf?__blob=publicationFile

[149] Schantz, Peter: Die Datenschutz-Grundverordnung - Beginn einer neuen Zeitrechnung im Datenschutzrecht. In NJW 2016, 1841 ff.

[150] Schneider, Jochen: Handbuch des EDV-Rechts. 4th Edition, Cologne, 2009.

[151] General Secretariat of the Council of the European Union: Manual of Precedents for Acts Established Within the Council of the European Union, 4th Ed., 2002. Available at: http://bookshop.europa.eu/en/manual-of-precedents-for-acts-established-within-the-council-of-the-european-union-pbQC4101381/

[152] Simitis, Spiros (ed.): Bundesdatenschutzgesetz Kommentar. 7th Edition, Baden-Baden, 2011.

[153] van der Sloot, Bart; Broeders, Dennis: Schrijvers, Erik (ed.): Exploring the Boundaries of Big Data. The Hague, 2016. Available at: http://www.ivir.nl/publicaties/download/1764

[154] Spiecker genannt Döhmann, Indra: A new framework for information markets: Google Spain. In Common Market Law Review 2015, 1033 ff.

[155] Spies, Axel: Cloud Computing: Keine personenbezogenen Daten bei Verschlüsselung. In MMR-Aktuell, 313727, 2011.

[156] Spindler, Gerald: Persönlichkeitsschutz im Internet - Anforderungen und Grenzen einer Regulierung. In Verhandlungen des 69. Deutschen Juristentages, Band I Gutachten, Munich, 2012.

[157] Spindler, Gerald: Persönlichkeitsrecht und Datenschutz im Internet - Anforderungen und Grenzen einer Regulierung. In NJW-Beilage, 98 ff., 2012.

[158] Spindler, Gerald: Datenschutz- und Persönlichkeitsrechte im Internet - der Rahmen für Forschungsaufgaben und Reformbedarf. In GRUR, 996 ff., 2013.

[159] Spindler, Gerald: Datenschutz- und Persönlichkeitsrechte im Internet - Der Rahmen für Forschungsaufgaben und Reformbedarf. In GRUR-Beilage, 101 ff., 2014.

[160] Spindler, Gerald: Durchbruch für ein Recht auf Vergessen(werden)? - die Entscheidung des EuGH in Sachen Google Spain und ihre Auswirkungen auf das Datenschutz- und Zivilrecht. In JZ 2014.

[161] Spindler, Gerald; Schuster, Fabian (eds.): Recht der elektronischen Medien. 3rd Edition, Munich 2015.

[162] Spindler, Gerald: Die neue EU-Datenschutz-Grundverordnung. In DB 2016, 937 ff.

[163] Spindler, Gerald: Selbstregulierung und Zertifizierungsverfahren nach der DS-GVO - Reichweite und Rechtsfolgen der genehmigten Verhaltensregeln. In ZD 2016, 407 ff.

[164] Spindler, Gerald; Schmechel, Philipp: Personal Data and Encryption in the European General Data Protection Regulation. In JIPITEC 2016, 163 ff. Available at: https://www.jipitec.eu/issues/jipitec-7-2-2016/4440/spindler_schmechel_gdpr_encryption_jipitec_7_2_2016_163.pdf

[165] Stadler, Thomas: Datenschutz: IP-Adressen als personenbezogene Daten. 27/06/2011. Available at: http://www.internet-law.de/2011/06/datenschutz-ip-adressen-als-personenbezogene-daten.html

[166] Sydow, Gernot; Kring, Markus: Die Datenschutzgrundverordnung zwischen Technikneutralität und Technikbezug - Konkurrierende Leitbilder für den europäischen Rechtsrahmen. In ZD, 271 ff., 2014.

[167] Taeger, Jürgen; Gabel, Detlev (ed.): Kommentar zum BDSG und zu den Datenschutzvorschriften des TKG und TMG. 2nd Edition, Frankfurt/Main, 2013.

[168] Tamó, Aurelia; George, Damian: Oblivion, Erasure and Forgetting in the Digital Age. JIPITEC 2014, 71 ff

[169] Tene, Omer: Privacy: The new generations. In International Data Privacy Law, 15 ff., 2011. Available at: http://idpl.oxfordjournals.org/content/1/1/15.full.pdf+html?sid=a11bd260-1434-4d5e-8e3b-c15b73574748

[170] Voigt, Paul: Datenschutz bei Google. In MMR, 377 ff., 2009.

[171] Walden, Ian: Anonymising Personal Data. In International Journal of Law and Information Technology 2002, 224 ff.

[172] Weichert, Thilo: Cloud Computing und Datenschutz. In DuD, 679 ff., 2010.

[173] Weichert, Thilo: EU-US-Privacy-Shield - Ist der transatlantische Datentransfer nun grundrechtskonform? In ZD 2016, 209 ff.

[174] Wieczorek, Mirko: Der räumliche Anwendungsbereich der EU-Datenschutz-Grundverordnung. In DuD, 644 ff., 2013.

[175] Wisskirchen, Gerlind: Grenzüberschreitender Transfer von Arbeitnehmerdaten. In CR, 862 ff., 2004

[176] Wolff, Amadeus; Brink, Stefan: Datenschutz in Bund und Ländern - Kommentar. Munich; 2013.

[177] Zanfir, Gabriela: The right to Data portability in the context of the EU data protection reform. In International Data Privacy Law 2012, 149 ff.

[178] Zuiderveen Borgesius, Frederik J.: Singling out people without knowing their names - Behavioural targeting, pseudonymous data, and the new Data Protection Regulation. In Computer Law & Security Review 2016, 256 ff.

Part II

Part II - Risk Assessment

Chapter 4

The methodology

4.1 Introduction

In this deliverable (D31.3) we revise and complete the description of the risk assessment methodology we introduced in D31.1 and D31.2. The ultimate goal of the *quantitative* methodology we propose is to provide an easy way to compute the residual risk remaining after the adoption of competing risk-reduction controls, supporting a quantitative cost-benefit calculation . We focus on process-related *leakage threats*, i.e., the disclosure of one or more information items to be exchanged in a multi-party protocol hosted on the cloud to participating parties who are not the originally intended recipients. The basic idea underlying our approach is using the *micro-economics* underlying the process to compute a quantitative estimate of the process' data leakage probability and impact. In a nutshell, our approach can be described as follows: we start by using the micro-economics underlying the business process to compute the *perceived unfairness* of the business process resource allocation. Then, we use this perceived unfairness to derive a continuous probability distribution associated to a formal random variable. Starting from the probability values of this distribution, we compute discrete probabilities of disclosure attacks associated to each subset of actors at each stage of the protocol, on the basis of the information items they could access up to that point. In previous deliverables we described the risk assessment for threats consisting mainly in disclosures due to individual agents. Here we describe the assessment of risk for threats posed by coalitions of actors, i.e.,, individuals or organisations participating in the business process or providers supplying the cloud infrastructure hosting the process. Firstly, we briefly recall the basics of our methodology.

4.2 Risk aware deployment of Secure Computation

Most risk management processes follow the same basic steps, although sometimes different terms are used to describe these steps. The following steps summarise management of risks regarding the enactment of business processes:

1. *Identification*. Uncover and describe risks that might affect the business process or its outcomes.

2. *Analysis*. Determine the likelihood and impact of each identified risk. You develop an understanding of the nature of the risk and its potential to affect the business process objectives.

3. *Evaluation*. Determine the risk magnitude, as an aggregation (usually the product) of likelihood and impact. Decide whether the risk is acceptable or whether it is serious enough to warrant treatment.

4. *Treatment*. Assess highly ranked risks and set out a plan to treat or alleviate them risks to achieve acceptable risk levels, e.g by installing appropriate security controls.

5. *Monitoring*. Monitor, track and review risks.

Our approach targets the risk that one or more participants could share among themselves or disclose to external parties the information they get to know by taking part in the process. Our technique consists of inserting an iterative sub-process as the fourth step of the risk management process. In this step the risk assessor can estimate the risks associated to misbehaviour of different subsets of the participants to the business process (henceforth called *actors*). This sub-process consists of a simulation where, step by step, the information exchanged among the actors is traced, and the collusion probability of each subset of actors is quantified, computing the overall likelihood of information disclosure. If the risk assessor decides that the risk level is not acceptable, she can explore alternative process enactments including the deployment of security controls like secure computation, run additional simulations and obtain a new risk analysis. At the end, a cost-benefit comparison will allow the risk assessor to decide if the cost of additional security controls is worth the reduction of the risk she has obtained in the new business process configuration.
Figure 4.1 shows the overall process.

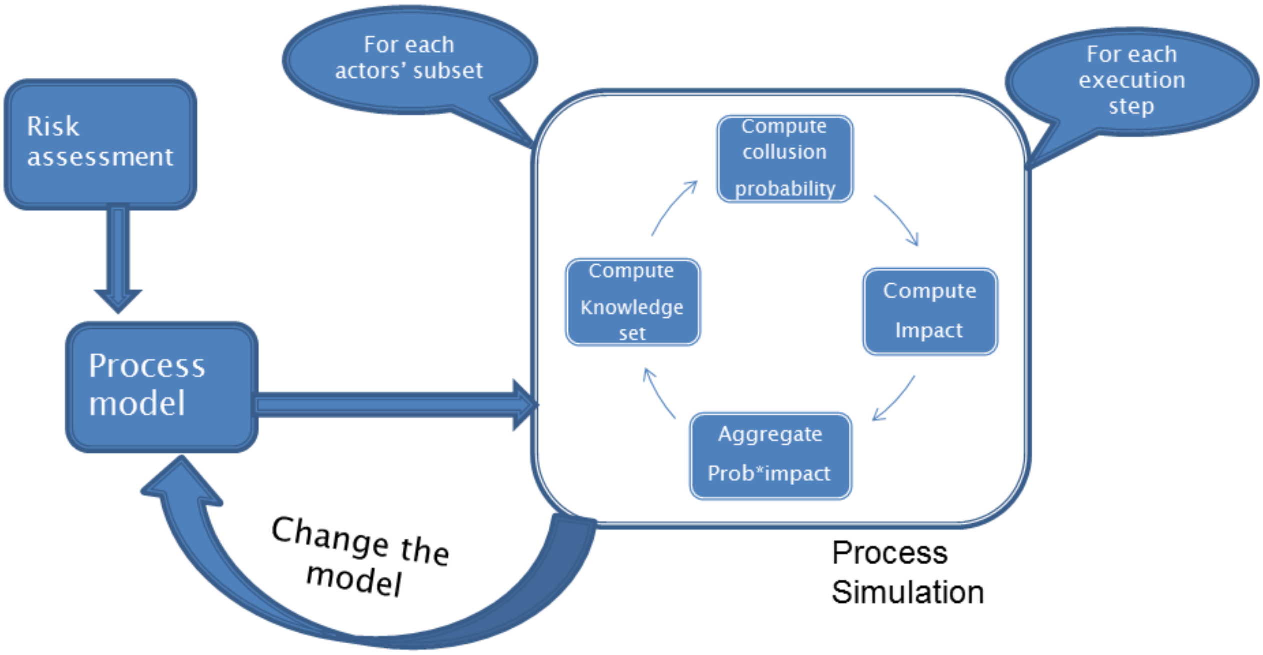


Figure 4.1: Our iterative process.

Let us now examine in detail our risk assessment methodology.

- We represent the business process P whose risk we want to assess via our *formalization of the business process model*, whose syntax is described in detail in section 4.4. Two types of actions are represented: (i) *message exchanges* and (ii) *local computations*. It is important to remark that our process model syntax expresses all possible execution paths *independently*, i.e., as separate models, providing a *process streamlining*, which includes *loop unrolling* and *re-encoding of conditions as parallel paths* [26].

- We compute (see again section 4.4), the probability of malicious behaviour on the part of each actor on the basis of a number of features, such as the unfairness of the redistribution of payoffs in the business process, (e.g. a benefit allocation structure that responds to organization's efficiency more than to fairness), the actor's greed, plus other context-related factors. At the same time, expert knowledge is used to estimate the impact of information disclosure on the business process outcome, quantifying the Value of Information (VoI) associated to each information item that could be disclosed.

Once the business process model has been designed, and the probabilities and impact values have been estimated, our iterative process starts by considering for each execution step of P and for each possible subset of actors:

- *reconstructible knowledge*, i.e *knowledge set* K_S for each subset $S \in 2^A$. The knowledge set includes all the knowledge that members of A can achieve by putting together the information they hold.
- *collusion probability* for each subset S in 2^A. Once again it is important to remark that this estimate needs to be process-specific (as it will take into account the micro-economics and social relations underlying P) and take into account multiple causes of collusion, including dysfunctional behavior, intervention of regulatory authority and others.
- *disclosure impact* of K_S for each subset S in 2^A.

Then, we compute the aggregation between (i) the collusion probability of each subset S in 2^A and (ii) the disclosure impact of K_S at each step of the process P, obtaining the *total risk* related to the process.

4.3 Modeling Security Controls

Our methodology aims to provide the risk assessor with the capability of comparing the risk alleviation due to alternative security controls used to secure the business process. For this purpose, it is possible to model the information exchange according to the security control adopted, and to automatically trace how the information actors can access (and therefore the risk profiles) are affected by the control's deployment.
On the basis of a comparative analysis among process enactments evolving different controls, the risk assessor can make the decision to trade-off cost for security. In this deliverable we focus on some specific techniques used within the PRACTICE project to perform secure computation (see also D31.1 for a more complete landscape of the available security controls). All the controls we consider belong to the family of *Secure Multiparty Computation* (SMC) techniques, which includes a wide range of security protocols, all supporting parties wishing to securely compute some agreed function on their private inputs. This notion includes special-purpose multi-party protocols supporting the evaluation of functions tailored to solve a specific problem such as *private set intersection* or *electronic voting*. It is important to remark that SMC protocols are added to ordinary business processes in order to achieve a security property like data confidentiality. For this reason, they can be seen as belonging to the category of security controls.

4.4 The Process Model

Let us now discuss our syntax for representing business process models. While similar syntaxes have been proposed for a number of purposes, from business process design to verification, our simple syntax is targeted to/for risk assessment purposes. We start by representing the business process' set of actors as a set $A = \{A_1, \ldots, A_n\}$. Each actor A_j holds a (possibly empty) information item $INFO_j$ whose content is used to generate messages to be exchanged during the business process' execution. Also, we denote by $\{I_{j,k}\}$ the impact of the disclosure of $INFO_j$ to A_k (as assessed by A_j). In principle, this impact can be positive or negative, and can depend on a number of factors, including the content of $INFO_j$ or of other information items. In our view, security controls (when present) are an integral part of the business process definition. In order to be able to represent all security controls of PRACTICE, message exchange in our process model is a general *timestamped choreography* [4] consisting of:

- *Messages*, i.e., triples (A_i, A_j, m_{ts}), where m_{ts} is (a part of) an $INFO$ item and ts is an integer representing a discrete time value[1]
- *Local computations* $(A_i, f(), INFO_{i,ts})$ i.e., functions computed by actors on (portions of) locally held information at a given time.

Next, we enrich our representation of business process actors and make it suitable for representing cloud-based computations. To this end, our actor set A becomes a (non-necessarily disjoint) triple $\{IN, COMP, RES\}$ where IN denotes actors holding non-empty information items (a.k.a. input nodes), while $COMP$ and RES are auxiliary sets of actors (a.k.a. *compute* and *result* actors) whose information items are initially empty. Such actors respectively perform local computations ($COMP$) and publish results (RES). The following constraints are in place for our cloud-enabled process models:

- *Separation of duties*: Sender actors belong to IN and $COMP$ only.
- *Local information integrity*: Any actor can send part of an $INFO$ item it holds entirely, or relay parts it has previously received from other actors.

Figure 4.2 shows a sample visual representation of a cloud-based process, where a buyer sends messages to two sellers who respond with their offers:

The Knowledge Set For each subset $S \in 2^A$, we can now compute the risk of disclosure for information shared within the actors belonging to S, at each time t. We proceed as follows: we consider all messages in the process incoming to actors belonging to S with timing $ts \leq t$. The (possibly empty) *common knowledge* of S, $K_S(t)$ is then composed of the $INFO$ items whose shares have been all received by members of S before time t, say $K_S(t) = \{INFO_{j_1}, \ldots, INFO_{j_h}\}$. The impact of the disclosure of this common knowledge on any actor $A_k \in A$ can be expressed in symbols as follows:

$$I_{S,k,t} = \sum_{p=1}^{h} I_{j_p,k} \tag{4.1}$$

and, in words, as the damage that members of S can do to A_k by disclosing all information items they can jointly reconstruct from the shares they hold at time t. Computing the risk posed by S to A_k also requires estimating the probability of members of S having colluded at time t (section 4.4). This risk can be written as follows:

$$R(A_k, E_S) = P_{S,t} I_{S,k,t} \tag{4.2}$$

Assuming that collusions happen independently, we can also write the total risk for A_k taking part in the process, as follows:

$$R(A_k, 2^A, \infty) = \sum_{S \in 2^A} I_{S,k,\infty} P_S \tag{4.3}$$

[1] For the sake of simplicity, in our model we assume synchronous clocks and instant message delivery.

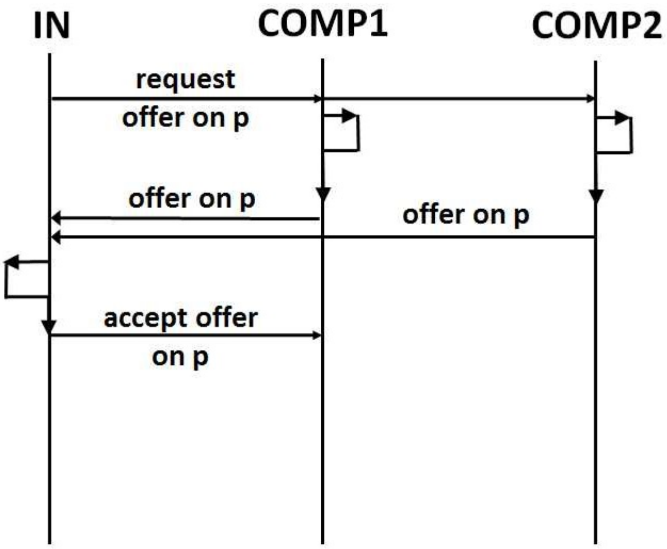


Figure 4.2: Visual representation of a sample cloud process model.

4.4.1 Reasoning about Shared Knowledge

A large amount of research work has been dedicated to understanding the process of reasoning about knowledge in a group and using this understanding for the analysis of complex systems [30, 13]. As regards the verification of cryptographic protocols modeled as sequences of encrypted messages exchanged to achieve some goals, *logic of knowledge* and *belief calculus* have been used extensively [8, 16].

Inference construction approaches try to use inference in some non-classical logics to establish required beliefs of the participants to a multi-party process. The *logic of authentication* introduced in [8], and commonly referred to as “BAN”, was one of the first successful attempts at representing and reasoning about the security properties of protocols, and has been extended and improved in successive works, as well as used as basis for the development of automatic verification tools [31, 21, 10]. In many cases, logics of knowledge have been used together with a formal model of a multi-agent system and a precise definition of a message trace [12, 20].

In this deliverable, we adopt a semi-formal approach, where we explicitly represent the knowledge sets, but do not equip our notation with the burden of a formal semantics. Our goal is simply to trace the evolution of the knowledge set associated to each subset of actors, and compute the overall risk by considering both the probability of dysfunctional behaviour on the part of the actors within the set and the value of the information they can hold or can reconstruct.

Securing processes Our methodology supports the risk assessor in constructing a representation of an unsecured business process, and then adding controls to secure it. Our syntax for expressing the operation of security controls over process messages starts from a plain-text exchange.

- Plaintext exchange $A \rightarrow A$
- Data Encryption $A \rightarrow E_{KA}[A]$
- Homomorphic Encryption $A \rightarrow HE_{KA}[A]$
- Secret sharing $A \rightarrow Sh_{1,n}[A] \oplus Sh_{2,n}[A] \oplus \cdots \oplus Sh_{n,n}[A]]$

4.4.2 The Role of the Cloud Provider

Under EU and US data protection laws (see part I), organisations remain responsible for the personal data of their customers and employees and must guarantee its security even when a third-party like a cloud provider processes the data on their behalf. The new EU General Data Protection regulation has been analysed in detail in the initial chapters of this deliverable. For this reason, data access issues have become a key consideration in any cloud provisioning arrangement. Data owners increasingly require outsourcing contracts to specify all details of cloud providers' security procedures, including prompt notification of any security breach and, above all, which techniques will be available to support cloud-based encryption/decryption services. Still, many cloud users have trouble in understanding and comparing available solutions.

When analysing the disclosure capability of subsets of actors taking part in a cloud-based business process, it is crucial to include a special participant, the *cloud provider*, since the messages exchanged among the actors as well as the information items stored in the actors' local memories could be potentially exposed to peeking on the part of the cloud provider and its employees. Of course, from the process model point of view the cloud provider may explicitly take part in the business process it hosts, i.e., be the source/destination of information exchanges, or simply provide the environment where such exchanges take place. We want to model this situation giving to the cloud provider the role of a potential *inside attacker*, who can access and/or manipulate the information items that are exchanged among the actors hosted on his platform. To this end, we enrich our notation for denoting process actors, by introducing the notion of actor "attached to" (i.e., running/using) services of a cloud provider. Each actor is labeled taking into consideration the hosting cloud. For example if C and D are two cloud providers, and A_1 and A_2 are two actors, the notation $A_1.C$ and $A_2.D$ means that the two actors run on different clouds, respectively managed by providers C and D. If also A_2 was running on the cloud provided by C, then the knowledge set of C will include all information items that A_1 and A_2 may observe. This situation is particularly relevant in the case of a secret-sharing-based security control; if secret information is shared between A_1 and A_2, the cloud provider can directly reconstruct the secret by peeking at the two shares. In notation: if at some time t the share $Sh_{1,2}[A]$ is held by A_1 and $Sh_{2,2}[A]$ is held by A_2 then C knows the entire information item A. This scenario can be modelled by what we call the *Cloud Transparency Equation* (CTE):

$$\cup_i KS(A_i.C) = KS(C) \tag{4.4}$$

4.4.2.1 Formalizing Cloud Transparency

In principle, process owners wishing to use the cloud for running their business processes could handle data protection by hosting controls (e.g., the ones performing encryption and decryption) on their own (or on a trusted third party's) premises. Unfortunately, a single untrusted cloud provider who is not allowed to decrypt customer data will be unable to deliver processing and display services. Using a different cloud for executing security controls may look a viable option; indeed PRACTICE itself deals with scenarios where multiple untrusted clouds and combinations of FHE, secret sharing and other Secure Multi-Party Computation (SMC) techniques are used to process encrypted data without decryption.

Generally speaking, however, our Cloud Transparency Equation (CTE) can be adapted to represent Cloud transparency, i.e., the effect of techniques for controlling access on the part of the cloud provider's *internal roles* to information exchanged by actors taking part in outsourced processes.

Most of these techniques (nested virtualisation, domain disaggregation and secure execution environment)[9, 19, 29, 5, 32], require architectural modifications, such as instrumenting one or more levels of the cloud stack with tamper-resistant hardware/software components aimed at protecting users' messages from unauthorised inspection and misuse on the part of the cloud provider's internal roles. For instance, in Infrastructure-as-a-Service (IaaS) cloud offerings, the employee of the cloud provider who administers the domain where the business process takes place can readily access the business process participants' messages and local memory. To represent the effect of architectural modification techniques, one can write in general that in a modified cloud stack the knowledge of the community of the actors will be larger that the one of the domain administrator role in charge of the process:

$$\cup_i KS(A_i.C) = K(C) > KS(Dom_C) \tag{4.5}$$

It is important to remark that our equational formalism can be used to specify which provider role has access to which information item. For instance, equation $KS(A.C) \in KS(Dom_C)$ specifies that the domain administrator has full monitoring of the information held by actor A, and similar equations can be written to denote that the domain administrator can observe a part of the messages and local memories involved in the process.

4.4.3 The Knowledge Transformation Rules

Our model includes a set of rules defining how the Knowledge Set evolves during the business process execution, and what are the information items that the actors can reconstruct during their participation to the business process. To this purpose, our notation includes a set of *knowledge transformation rules*, that can be used to automatically compute the information items shared among a given set of actors, as resulting from the previously exchanged messages and a-priori knowledge. The rule explicitly represent basic facts, such as the possibility of reading the content of an encrypted message when holding both the key and the message itself, or of reconstructing a shared secret when holding all its shares. Such rules are necessary to automatically reason on shared knowledge and are the basis of the operation of our software tool, described in detail in section 6.

The Encryption rule This rule expresses the fact that holding the decryption key together with holding a data item encrypted with it is equivalent to holding the plaintext.

$$K_A \oplus E_{K_A}[X] \rightarrow X \tag{4.6}$$

The Share Reconstruction rule This rule expresses the fact that holding all the shares of a data item is equivalent to holding its plaintext.

$$Sh_{1,n}[A] \oplus Sh_{2,n}[A] \oplus \cdots \oplus Sh_{n,n}[A] \rightarrow A \tag{4.7}$$

The Homomorphic Encryption rule This rule expresses the fact that holding operands homomorphically encrypted w.r.t. an arithmetical operation is equivalent to holding the encrypted result of the operation.

$$HE_{K_A}[X] \oplus HE_{K_A}[Y] \rightarrow HE_{K_A}[X \otimes Y] \tag{4.8}$$

These equational annotations can be developed to express the properties of a number of different security controls applicable to the business processes. We will further discuss this notion in the next chapter.

Chapter 5

Handling Coalition Threats

Our framework addresses the management of threats from insiders in a cloud based collaborative business process. Attacks can consist in information leakage/exfiltration, information tampering, misrepresentation of information within the process or in general any (non agreed upon) deviation from the collaborative protocol. We do so by analyzing some major factors which can motivate an attack, assessing the risk associated to individual factors and indicating the countermeasures that can mitigate the risk.

Several sources [17, 23, 6] acknowledge that the motivations for attacks from insiders employees or organizations can belong mainly to the following categories: economical/financial, competitive advantage, political, for revenge, curiosity or fun, power, or peer recognition.
In our setting, only the financial and competitive advantage motivations apply to economical organizations, furthermore – given the assumptions we adopt w.r.t. the information available to the risk assessment process – the other motivations cannot be investigated (for instance information on the political orientation or on the desire of peer recognition by an employee is typically not available). Therefore, we only deal with quantifiable factors of economical character.

We identified two main factors in this category: lack of (economical) satisfaction and (economical) greed. Both can apply to individuals or to organisations and can be assessed on the basis of quantitative data and quantitative estimates from domain experts. We quantify the strength of those factors by means of coalition analysis of the collaborative system.

5.1 Assessing Risk for coalitions of actors

We start by introducing some key concepts of game theory we will need for quantifying the likelihood of threats. These concepts, discussed in detail in Section 5.2 assume that the micro-economics underlying a business process can be modeled as a coalition game where players include all actors participating in the process in various roles. In order to uniform our terminology to the one used in the game-theoretical literature, in the following we will refer to the set of all business process participants as *grand coalition*. Indeed, it is known that grand coalition instability can lead to the formation of sub-coalitions whose interests are not aligned with those of the grand coalition. In the same section, we will recall the concept of fair division and its implications from the sub-coalitions point of view. After establishing the terminology and some background, in section (Section 5.3) we describe the likelihood computation, Then, in Section 5.4 we show a complete use case.

5.2 Coalition analysis concepts: stability and fairness

Coalitional (or cooperative) games are a branch of game theory in which one can model cooperation or collaboration among agents. In a coalitional game, we focus on what groups of players can achieve rather than on what individual players can do. Intuitively, stability of a coalitional game means that the game outcome is immune to deviations by groups of players, i.e.,, no subset of players can unilaterally improve their outcome.

5.2.1 Coalitional games with transferrable utility (TU games)

If we have a set $\mathscr{N}$ containing $N = |\mathscr{N}|$ agents, a coalition is defined as follows.

Definition 1 (Coalition) *A* **coalition** *C is a set of agents: $C \in 2^{\mathscr{N}}$.*

Two main classes of games are considered in Coalitional Game Theory: Transferrable Utility Games and Non-Transferrable Utility Games. A Transferrable Utility Game is a game where

- two agents can **compare** their utility
- two agents can **transfer** some utility

A Non Transferrable Utility Game is a game where

- it is not always possible to compare the utility of two agents
- it is not always possible to transfer the utility between two agents.

We focus on Transferable Utility (TU) games. In a TU game, it is assumed that the earnings of a coalition can be expressed by one number. One may think of this number as an amount of value generated by the process, which can be distributed among the actors in any conceivable way - including negative payments - if the grand coalition is actually formed. In general terms, this number is an *amount of utility* and our assumption in the following is twofold: (i) individual utilities of the business process participants can be expressed in monetary terms (ii) it makes sense to transfer/share this utility among the participants.

Definition 2 (Valuation) *A* **valuation function** *μ (or* **characteristic function***) is a set function that associates a real number $\mu(C)$ to any subset of a set of agents, i.e.,*

$$\mu : \; 2^{\mathscr{N}} \rightarrow \mathbb{R}$$

Definition 3 (TU game) *A* **TU game** *is a pair $(\mathscr{N}, \mu)$, where $\mathscr{N}$ is a set of agents and μ is a valuation function.*

We are especially interested in games where the collaboration produces a surplus (in a sense they are the non-trivial collaborative games). This idea is captured by the concept of *essential games*.

Definition 4 (Essential game) *A game $(\mathscr{N}, \mu)$ is essential if $\mu(\mathscr{N}) > \sum_{i \in \mathscr{N}} \mu(i)$.*

In an essential game there is a positive difference between the minimal values that each player can attain individually and the total value that can be attained by the whole grand coalition players: in a TU game this extra wealth is exactly the utility that can be allocated among the players.

5.2.2 Issues in TU games

In coalitional game theory the following two main issues regarding TU games are addressed:

- what coalitions will form or which, once formed, will be stable?
- how to reward each member of a coalition when a task is completed?

5.2.3 Definition and properties of valuations

In algebra (in particular in algebraic geometry or algebraic number theory), a valuation is a function on a field that provides a measure of size or multiplicity of elements of the field. Here, we will consider valuations over the lattice of subsets of a given set, a.k.a.its boolean or powerset algebra.

5.2.3.1 Measures or capacities

Definition 5 (Measure or Capacity) *A* measure *is a map from the powerset algebra of a set $\mathscr{N}$ to a scalar field, in our case $\mathbb{R}$,*

$$\mu : C \in 2^{\mathscr{N}} \to \mathbb{R}$$

vanishing at the empty set, i.e., such that

1) $\mu(\emptyset) = 0$ *(0-normalization)*

Since no *additivity* or *positivity* is implied, this map is more soundly called *non-additive signed measure*, or **capacity**.

It is important to remark that a **valuation** in TU game theory is an example of such a measure. Every measure can be taken to represent the characteristic function of a game, thus defines univocally a game.
In our application field we are mostly interested in *non-negative measures*.

Definition 6 (Non-negativity) *A measure is said* non-negative *if*

2) $\mu(S) \geq 0 \qquad \forall S \in 2^N$ ***(non-negativity)***

hereafter, when using the term *measure* or *capacity*, will always imply the non-negativity condition.

Definition 7 (Normalization) *A measure is said* normalized*, or 1-normalized, if*

3) $\mu(N) = 1$ ***(normalization)***

If a measure is both 0-normalized and 1-normalized, it is said $(0,1)$-normalized. For instance, standard probability measures are $(0,1)$-normalized.
We are now ready for introducing a simplifying notion, i.e., the one of $(0,1)$-normalized games.
For essential games, it is always possible to transform a game $(\mathscr{N},\mu)$ from the non-normalized form to a $(0,1)$-normalized form, through the following transformation.

Proposition 1 ($(0,1)$-Normalization of an essential game) *If $(\mathscr{N},\mu)$ is an essential game, then it is equivalent to a $(0,1)$-normal game $(\mathscr{N},\mu^*)$ defined by*

$$\mu^*(C) = \frac{\mu(C) - \sum_{i\in C}\mu(i)}{\mu(\mathscr{N}) - \sum_{j\in\mathscr{N}}\mu(j)}$$

5.2.3.2 Monotonicity

Definition 8 (Monotonicity) *A measure is said to be increasing (decreasing)* monotone *if the measure of a set cannot be smaller (larger) than the measure of one of its subsets.*

4.a) $S \subseteq T \Rightarrow \mu(S) \leq \mu(T)$ *(**increasing monotonicity**)*

4.b) $S \subseteq T \Rightarrow \mu(S) \geq \mu(T)$ *(**decreasing monotonicity**)*

Strict monotonicity has the signs $<$ *and* $>$*.*

A monotonic measure can be used as the seed for a parametric family of measures satisfying probabilistic inequalities. For instance:

Example 1 (Monotonic measures)

- *the **max** measure* $\mu(S) \equiv \max_{i \in S}(\mu(i))$ *is monotonically increasing*
- *the **min** measure* $\mu(S) \equiv \min_{i \in S}(\mu(i))$ *is monotonically decreasing*
- *Probability Measures are monotonically increasing*

For monotonically increasing measures the following inequalities hold

- $\mu(A \cup B) \geq \max(\mu(A), \mu(B))$, since $A \cup B$ is a superset of both A and B
- $\mu(A \cap B) \geq \min(\mu(A), \mu(B))$, since $A \cap B$ is a subset of both A and B

Definition 9 (Capacity) *A capacity is a measure holding the following properties:*

- $(0,1)$*-normalized*
- *monotone*

Example 2 (Capacities) *For instance*

- *The "Glove Game" defined by*
  - $\mu(\emptyset) = \mu(1) = \mu(2) = \mu(3) = \mu(12) = 0,$
  - $\mu(13) = \mu(23) = \mu(123) = 1$
- *The "Airport Runway Game" defined by*
  - $\mu(\emptyset) = 0, \mu(1) = 8/18,\ \mu(2) = 11/18,\ \mu(3) = 13/18,\ \mu(4) = 18/18$
  - *and* $\mu(S) = \max_{i \in S} \mu(i)$

5.2.3.3 Additivity, non-additivity, super- and sub-additivity

Definition 10 (Additivity) *A measure is said to be* additive *if, for every disjoint pair of sets S, T, it maps their union into the sum of their individual measures, i.e., if*

5) $\mu(S \cup T) = \mu(S) + \mu(T) \qquad \forall S, T \text{ s.t. } S \cap T = \emptyset$ *(**additivity**)*

A measure is said to be non-additive if *at least for a pair of subsets* the above condition does not hold.

Example 3 (Additive and non-additive measures) *For instance*

- *Let a measure count the number of pairs in a set $\mu(A) = |A|(|A|-1)$, then it is non-additive.*
- *The classical probability measures are examples of additive measures.*

If a coalitional game is defined by an additive measure, then it is said additive or **inessential**, since it is trivial from the game theoretic point of view (in an inessential game forming a coalition does not bring any benefit to the participants).
It is easy to see that **additivity** implies **monotonicity**, but not the other way round. Non-additive monotonic measures generalise the additive ones.

The requirement of additivity (countable or finite) of classical measures is based on the assumption that disjoint sets are **non-interactive** with respect to the measured property. This assumption is too restrictive in some application contexts.
A monotone measure is able to capture the following situations relating to two disjoint sets A and B:

a) $\mu(A \cup B) > \mu(A) + \mu(B)$ (positive or cooperative interaction, synergy)

b) $\mu(A \cup B) = \mu(A) + \mu(B)$ (non-interactivity of A and B)

c) $\mu(A \cup B) < \mu(A) + \mu(B)$ (negative or inhibitory interaction)

An additive theory, such as probability theory, based on classical measure theory, is capable of capturing only situation (b). The theory of monotone measures provides a considerably broader framework. The wide range of application of non-additive measures is due to the fact that in any situation with more than minimal complexity the effect of the ensemble of parts is different from the sum of the effects of the parts taken individually.
Notice that in a specific setting there could be at the same time positive interactions between some sets and negative interactions between other sets. They allow us to represent interaction among the distinct elements.
For example, we might have

- $\mu(A \cup B) > \mu(A) + \mu(B)$ (positive interaction between A and B) and
- $\mu(A \cup C) < \mu(A) + \mu(C)$ (negative interaction between A and C).

Nonetheless, in some remarkable situations the interaction is always cooperative, while in other remarkable situations the interaction is always inhibitory.

Definition 11 (Super-additivity) *A measure is said* ***super-additive*** *if for all the pairs of disjoint sets the measure of the union is not smaller than the sum of the measures of the individual sets:*

6) $\mu(S \cup T) \geq \mu(S) + \mu(T) \quad \forall S, T$ *s.t.* $S \cap T = \emptyset$ ***(super-additivity)***

The negative of a super-additive function is a sub-additive function.

Example 4 (Super-additive measures) *For instance*

- *A measure counting the number of pairs in a set* $\mu(A) = |A|(|A|-1)$.
- *The Glove Game measure is super-additive.*

Definition 12 *A measure is said* ***sub-additive*** *if for all the pairs of disjoint sets the measure of the union is not greater than the sum of the measures of the individual sets:*

7) $\mu(S \cup T) \leq \mu(S) + \mu(T) \quad \forall S, T$ *s.t.* $S \cap T = \emptyset$ ***(sub-additivity)***

Example 5 (Subadditive measures) *For instance*

- *The* min *function:* $\mu(S \cup T) \equiv \min(\mu(S), \mu(T))$.
- *The* max *function:* $\mu(S \cup T) \equiv \max(\mu(S), \mu(T))$.
- *The Airport Runway Game is sub-additive.*

The negative of a sub-additive function is a super-additive function.

Special classes of measures Sub-additive measures and super-additive measures are two special types of monotone measures. Other special types include classical, additive measures (i.e., probability measures), the classical (crisp) possibility and necessity measures. Each special type of monotone measures can be used for formalising a certain type of uncertainty. A remarkable type of uncertainty can be formalised based on a property stronger than super-additivity: modularity.

5.2.3.4 Modularity, super- and sub-modularity

While additivity/non-additivity represent the behaviour of a measure with respect to the union of **disjoint sets**,*modularity*, refers to the union of **any pairs of sets**.

Definition 13 (Modularity) *A measure is said to be* modular *if the following inclusion-exclusion equation holds for any pairs of sets*

- $\mu(S \cup T) = \mu(S) + \mu(T) - \mu(S \cap T) \qquad \forall S, T \in 2^{\mathcal{N}}$ *(modularity)*

It is easy to see that if a measure is **additive** it is also **modular**. We are now ready to define super-modularity

Definition 14 (Super-modularity of order 2, a.k.a. convexity, a.k.a. 2-monotonicity) *A measure is said super-modular if in the inclusion-exclusion relation the* $=$ *sign is replaced by* $\geq$*:*

- $\mu(A \cup B) \geq \mu(A) + \mu(B) - \mu(A \cap B) \qquad \forall A, B \in 2^N$ *(super-modularity)*

Since it involves pairs of sets it is also called *super-modularity of order* 2.

- For monotone measures, super-modularity is a condition stronger than super-additivity, i.e., super-modular measures are monotone non-decreasing. Super-modular measures of order 2 are also called 2-monotone measures.

- Super-modularity, captures the *increasing returns* property: for any set $X \subseteq Y \subseteq \mathscr{N}$, and any $i \in \mathscr{N}$, the inequality
$$\mu(Y \cup i) - \mu(Y) \ \geq \ \mu(X \cup i) - \mu(X)$$
expresses the fact that the marginal contribution of an element i to a set is a non-decreasing function of the set. Rearranging the definition one gets $\mu(Y \cup i) \ \geq \ \mu(Y) + \mu(X \cup i) - \mu(X)$, i.e the previous expression if one takes $A = Y$ and $B = (X \cup i)$.

- In Game Theory, the super-modularity of a game (a.k.a. *convexity*), captures the intuitive notion that in the game incentives for joining a coalition increase as the coalition grows.

Definition 15 (Sub-modularity (of order 2)) *A measure is said super-modular if in the inclusion-exclusion relation the $=$ sign is replaced by $\leq$:*

- $\mu(S \cup T) \leq \mu(S) + \mu(T) - \mu(S \cap T) \qquad \forall S, T \in 2^N \qquad$ *(sub-modularity)*

- For monotone measures, sub-modularity is a condition stronger than sub-additivity, i.e., sub-modular measures are monotone non-increasing.

- Sub-modularity, can be used to capture the *diminishing returns* property: for any set $A \subseteq B \subseteq N$, and any $i \in N$, the inequality
$$\mu(B \cup i) - \mu(B) \ \leq \ \mu(A \cup i) - \mu(A)$$
expresses the fact that the marginal value of an element i to a set A is a non-increasing function of the set A.

- In coalitional Game Theory, the sub-modularity of a game (a.k.a. *concavity*), captures the intuitive notion that in the game incentives for joining a coalition decrease as the coalition grows.

This property can be easily generalized to higher orders. For instance, super-modularity of order 3 is defined as follows.

Definition 16 (Super-modularity of order 3, a.k.a. 3-monotonicity)

$$\begin{aligned}\mu(A \cup B \cup C) \ \geq \ & \mu(A) + \mu(B) + \mu(C) \\ & -\mu(A \cap B) - \mu(A \cap C) - \mu(B \cap C) \\ & +\mu(A \cap B \cap C) \qquad\qquad \forall A, B, C \in 2^{\mathscr{N}}\end{aligned}$$

Every super-modular measure of order 3 is also super-modular of order 2. The above expression can be rewritten more compactly by considering the families of 3 sets and denoting the members of each family by $A_1, A_2, A_3 \in 2^{\mathscr{N}}$

$$\mu\left(\bigcup_{i=1}^{3} A_i\right) \geq \sum_{\emptyset \neq K \subset \{1,2,3\}} (-1)^{(|I|+1)} \mu\left(\bigcap_{i \in I} A_i\right) \qquad A_1, A_2, A_3 \in 2^{\mathscr{N}}$$

where I is a set of indexes.
The corresponding order 3 sub-modularity can be defined through the above expressions, by changing $\geq$ into $\leq$.
All the property described above can be generalized to any order $k \leq n$ by considering the families of k sets and denoting the members of each family by $A_1, A_2, \ldots, A_k \in 2^N$. For the sake of brevity we consider only super-modularity, the corresponding sub-modularity definitions can be obtained by changing $\geq$ into $\leq$.

Definition 17 (Super-modularity of order k, a.k.a. k-monotonicity) *A non-additive measure is said to be order k super-modular if for all collections of k subsets the following relation holds*

$$\mu\left(\bigcup_{i=1}^{k} A_i\right) \geq \sum_{\emptyset \neq I \subset \{1,2,\ldots,k\}} (-1)^{(|I|+1)} \mu\left(\bigcap_{i \in I} A_i\right) \qquad A_1, A_2, \ldots, A_k \in 2^N$$

where I is a set of indexes.

Every k-monotone measure of order $k > 2$ is also of order $k' \leq k$, but, a capacity of order k is not necessarily a measure of any higher order of monotonicity.

Definition 18 (Totally monotone measures) *In a finite universe of size $N = |\mathcal{N}|$ a k-monotone measure with $k = N$ is said* totally monotone *(or ∞-monotone).*

5.2.3.5 Choquet Capacities

The capacity of a set is a measure of that set's "size". Unlike, say, Lebesgue measure, which measures a set's volume or physical extent, capacity is a mathematical representation of a set's ability to hold a measurable content. The notion of capacity of a set was introduced by Gustave Choquet in 1950.

Definition 19 (Order 2 Choquet capacity) *A non-additive measure which is super-modular (*2*-monotone) is said to be an order* 2 *Choquet capacity (or* 2*-monotone Choquet capacity).*

Definition 20 (Order k Choquet capacity) *An order k super-modular (k-monotone) non-additive measure is said to be an order k Choquet capacity (or k-monotone Choquet capacity).*

For convenience, monotone measures that are not required to satisfy the above equation are viewed as Choquet capacities of order 1 (i.e., 1-monotone Choquet capacities).

Definition 21 (Totally monotone Choquet capacities) *Order $N = |\mathcal{N}|$ Choquet capacities are also called* totally monotone *capacities.*

- Additive capacities are clearly totally monotone.
- Every Choquet capacity of order $k > 2$ is also of order $k' = 2, 3, \ldots, k$, but, a capacity of order k is not necessarily a capacity of any higher order.
- Hence, the Choquet capacities of order 2 (2-monotones) are the most general.
- Those which are of order k $\forall k \geq 2$ are called Choquet capacities of infinite order (or ∞-monotones)
- belief functions of the Dempster-Shafer theory are totally monotone Choquet capacities

- each order k of Choquet capacity can be used to formalize a different theory of uncertainty
- probability measures are special totally monotone Choquet capacities for which all the inequalities in the modularity expression collapse to equalities.

By the above definitions derives the following nesting among classes:

- additive capacities are totally monotones,
- totally monotone capacities are super-modular (convex)
- and super-modular capacities are super-additive.

5.2.4 Coalition stability concepts

We are now ready to study allocations that determine stable coalitions.
Specifically, let the characteristic function μ of a game (the $(2^N - 1)$ *valuations* for all the possible non-empty coalitions) be given. Consider for the sake of simplicity the grand coalition $\mathscr{N}$. We adopt the convention that the valuation of such coalition is equal to one: $\mu(\mathscr{N}) = 1$. This value represents the total returns earned by the coalition in this specific game.
One can assign to each player $i \in \mathscr{N}$ of the gran coalition as a payoff a share x_i of the total returns, and ask if the proposed solution has determinate properties.

Definition 22 (Feasibility) *A solution is feasible if it does not exceed the total worth of the grand coalition:*

$$\sum_{i\in\mathscr{N}} x_i \leq \mu(\mathscr{N})$$

Definition 23 (Efficiency) *A solution is efficient if the payoff distribution is an allocation of the entire worth of the grand coalition to all agents:*

$$\sum_{i\in\mathscr{N}} x_i = \mu(\mathscr{N})$$

Definition 24 (Anonymity) *A solution is independent of the names of the players.*

Definition 25 (Individual rationality) *A solution is individually rational if each agent obtains at least its self-value as a payoff.*

$$\forall i \in \mathscr{N} \qquad x(i) \geq \mu(\{i\})$$

Definition 26 (Imputation) *An imputation is a payoff distribution x that is*

- efficient
- *and* individually rational

Definition 27 (Coalitional rationality/group rationality) *A solution is group-wise rational if each group obtains at least its self-value as the sum of its component's payoffs*

$$\forall C \subseteq \mathscr{N} \qquad \sum_{i\in C} x_i \geq \mu(C)$$

Indeed, on one hand, if $\sum_{i\in C} x_i < \mu(C)$, no group would be ready to accept.
Group rationality ensures individual rationality as a special case.
Occasionally, hereafter, we will use the shorthand notation $x(C) \equiv \sum_{i\in C} x_i$: in those terms efficiency is expressed by the condition $x(N) = \mu(N)$, while group rationality is expressed by the condition $x(C) \geq \mu(C), \ \forall C \subseteq \mathscr{N}$.

5.2.4.1 The Core

Let us consider a TU game $(\mathscr{N},\mu)$ and its grand coalition. The notion of Core, first introduced by Gillies [15], is a natural way to define stability: a payoff distribution lies in the Core when no sub-group of agents has any incentive to form a different coalition.

Definition 28 (Core) *The core is the set of payoff allocations satisfying* efficiency *and* coalitional rationality.

$$Core(\mathscr{N},\mu) \equiv \left\{ x \in \mathbb{R}^N \;\middle|\; \left(\sum_{i\in\mathscr{N}} x_i = \mu(\mathscr{N}) \right) \wedge \left(\forall C \subseteq \mathscr{N} \qquad \sum_{i\in C} x_i \geq \mu(C) \right) \right\}$$

Equivalently one can say that the core is the collection of group-rational imputations.

Example 6 (Core of some 2-player games) *For instance*

- *Let us consider the following two-player game* $(\{1,2\},\mu)$ *where* $\mu(1)=5$*,* $\mu(2)=5$*, and* $\mu(1,2)=20$*. The core of the game is a segment defined as follows:*

  $$Core(\{1,2\},\mu) = \{(x_1,x_2)\in\mathbb{R}^2 | x_1 \geq 5, x_2 \geq 5, x_1+x_2=20\}.$$

  Notice incidentally that the segment contains the point $(\frac{1}{2},\frac{1}{2})$*: due to the symmetry of the two players' contributions this can be taken as a fair allocation, while all the other allocations in the core are, to some degree, unfair.*

- *Let us consider the "glove game"* $(\{1,2,3\},\mu)$ *where* $\mu(1)=\mu(2)=\mu(3)=0$*,* $\mu(1,2)=\mu(1,3)=1$*,* $\mu(2,3)=0$ *and* $\mu(1,2,3)=1$ *. The core of the game is defined by:*

  $$Core(\{1,2,3\},\mu) = \{(x_1,x_2,x_3)\in \left(\mathbb{R}^+\right)^3 | \, x_1+x_2 \geq 1, x_1+x_3 \geq 1, x_1+x_2+x_3 = 1\} = (1,0,0).$$

  Here the core is a single point. Notice incidentally that the corresponding allocation is not fair.

The core condition for coalition stability is so strong (the condition $\sum_{i\in C} x_i \geq \mu(C)$ has to be true for all the subsets C of $\mathscr{N}$), that some games may have an empty core: not all the players and groups of players can be satisfied simultaneously. The Core is empty when at least one player is dissatisfied by the payoff allocation. Such player can be said to "block" the coalition; in other words she can rise an objection against the formation of such coalition, or threaten to leave the coalition if the coalition is already in place. When focusing on the grand coalition, the core is the set of allocations such that the grand coalition cannot be blocked.

The literature on TU games provides several sufficient conditions for the the core of a game to be non-empty, or for the core of a game to be empty (for a summary see for instance Owen [24] or Peleg and Sudhölter [25]). For instance the core of inessential games, i.e., those games such that $\sum_{i\in\mathscr{N}} \mu(i) > \mu(\mathscr{N})$, is trivially empty; on the other hand convex games have a non empty core. The Bondareva and Shapley Theorem provides a characterization of the non-empty core games using the concept of balancedness of the game (see [25]).

Checking if a payoff allocation lies in the core. However the issue of core non-emptiness falls outside the scope of the present document: here we are interested not in characterizing a game's core, but in assessing the stability of individual payoff allocations. To this purpose one has to check that the conditions of efficiency and group rationality are fulfilled.
Checking efficiency amounts to checking the equality $(\sum_{i\in\mathcal{N}} x_i = \mu(\mathcal{N}))$.
The basic algorithm for checking the inequality $(\sum_{i\in C} x_i \geq \mu(C))$ for all the possible subsets $C \subseteq \mathcal{N}$, thus it is exponential in the number N of players, however a polynomial time algorithm exists for performing the check: indeed, a vertexes of the convex region representing the core can be found in polynomial time; once all the vertexes are found it is easy to check if a point is inside the region.[1].
In special conditions this check can be performed in polynomial time. It is important to remark that in the context of our application to the study of coalitions underlying cloud-based business processes not all coalitions are possible: so, the number of checks to be performed is considerably reduced, so as that even the basic algorithm might be affordable.

5.2.4.2 Augmentations of the Core

The concept of Core has been declined in a few variants, some addressing the problem that the Core can be empty. A possible objection to the definition of Core is that sometimes a member of a coalition might be seen as potentially blocking, while in practice objecting to the coalition formation or threatening to leave would bring her a very small gain. If it is unlikely that an actor blocks a coalition for a small utility gain. This considered, one can establish a sort of insensitivity threshold ε, i.e., an amount that is not considered by an actor as a interesting gain: this leads to the concept of *ε-core*. The ε-core comes into two variants: the *weak ε-core* and the *strong ε-core*, that differ by the fact that the tolerance is encoded at per player player level or at coalition level. The standard way of formalizing these concepts is the following.

Definition 29 (Strong ε-core) *The strong ε-core is the set of the efficient allocations that for every $C \subseteq \mathcal{N}$ satisfy*

$$\sum_{i\in C} x_i \geq \left(\ \mu(C) - \varepsilon\ \right)$$

Definition 30 (Weak ε-core) *The weak ε-core is the set of the efficient allocations that for every $C \subseteq \mathcal{N}$ satisfy*

$$\sum_{i\in C} x_i \geq \left(\ \mu(C) - |C|\cdot\varepsilon\ \right)$$

The weak version of the concept takes into account the size of the coalition in defining the tolerance. This concept can be used to solve the empty core issue: if one increases the value of ε, the ε-core at some point becomes non-empty: for larger values of ε the core is non-empty. The least value of ε for which the core is non-empty is taken as special value of the game and the corresponding ε-core is called *least ε-core*.

Even if in our application we do not focus on the issue of core emptiness, we briefly consider the ε-core concept because it is important for a more realistic modeling of stability problems.

[1] A vertex of the core can be found using the following algorithm. Let $\pi : N \to N$ be a permutation of the players, and let $S_i = \{j \in N : \pi(j) \leq i\}$ be the set of players ordered 1 through i in π, for any $i = 0,\ldots,N$, with $S_0 = \emptyset$. Then, the payoff $x \in \mathbb{R}^N$, defined by $x_i = \mu(S_{\pi(i)}) - \mu(S_{\pi(i)-1}), \forall\, i \in N$ is a vertex of the core of μ. Any vertex of the core can be constructed in this way by choosing an appropriate permutation π.

5.2.4.3 Games with coalitional structure

So far we focused on the formation of the grand coalition, i.e., the one formed by all participants. This corresponds to the practical situation in which the super-additivity of the valuation function is either explicitly stated or implicitly assumed. However, when the valuation function is not super-additive, players may have an incentive to form a different partition. Furthermore sometimes some coalitions are excluded by necessity (e.g. impossibility to measure or communicate) or due to an exogenous choice.
To model these cases one has to consider *coalitional structures* (CS): the concept was introduced by Auman and Dreze [3]. A coalitional structure S is a partition of the grand coalition: S is a CS if $S = \{C_1, C_2, \ldots, m\}$, with $\cup_{k=1}^{m} C_k = \mathscr{N}$ and $\forall i \neq j \; C_i \cap C_j = \emptyset$.

Definition 31 (Game with coalitional structure) *A game with coalitional structure is a triplet $(\mathscr{N}, \mu, S)$, where $(\mathscr{N}, \mu)$ is a TU game and S is a specific CS. In addition, transfer of utility is allowed only within coalitions of S and not between coalitions of S.*

Notice that the problems of deciding which coalition to form and how to share the coalition's revenue so as to achieve stability are independent and decoupled. Also in the previous section the CS was fixed: it consisted in the trivial partition defined by the grand coalition. As before, here we address the problem of the stability of individual coalitions, by examining the payoff distributions within each coalition and checking efficiency and group rationality conditions.

Definition 32 (Feasible payoff) *Let $(\mathscr{N}, \mu, S)$, be a TU game with CS. The set of feasible payoff distributions is*

$$X_{(\mathscr{N}, \mu, S)} = \{x \in \mathbb{R}^N | \forall C \in S, x(C) \leq \mu(C)\}$$

In analogy to the previous definition of efficiency (where the condition was $x(N) = \mu(N)$), but restricted to the coalitions allowed by the CS we define the following.

Definition 33 (Efficiency w.r.t. a CS) *A payoff distribution x is efficient with respect to a CS S when*

$$\forall C \in S \qquad x(C) = \mu(C)$$

In the context of CS, a payoff distribution is an imputation if it is efficient w.r.t. the current CS and it fulfills individual rationality (i.e., $\forall i \in \mathscr{N}, x_i \geq \mu(i)$). We denote set of all the imputations for a CS by $Imp(S)$. The definition of the core in TU games with CS is the following.

Definition 34 (Core of a game with CS) *Let $(\mathscr{N}, \mu, S)$, be a TU game with CS. The core of the game is the set of imputations $x \in Imp(S)$ that are group rational for the coalitions of the CS.*

5.2.4.4 The nucleolus

The concept of nucleolus, was introduced by Schmeidler [27] and it is based on the notion of a coalition discontent or dissatisfaction, measured through a quantity called *excess*.

Definition 35 (Excess) *Let $(\mathscr{N}, \mu)$ be a TU game, $C \subseteq \mathscr{N}$ be a coalition and x be a payoff distribution over $\mathscr{N}$. The* **excess** *$e(C, x)$ of a coalition C at x is the quantity*

$$e(C, x) \equiv \mu(C) - x(C)$$

When the excess is positive, the coalition is not receiving all its worth, and might be unsatisfied and possibly misbehave. Note that if a payoff distribution lies in the core there cannot be any complaint, because the core constraint of group rationality is equivalent to non-positive excess.

The nucleolus, by using the concept of excess, is a notion especially useful in dealing with payoff distributions which are out of the core. The idea behind the nucleolus notion is that sometimes a coalition (or a coalitional structure) is forced to live with a level of dissatisfaction (by necessity or by external choice): the nucleolus corresponds to the configurations which, given the total worth of the grand coalition, minimize the dissatisfaction.

Minimization implies comparisons (using an evaluation function): in this case we need to perform pairwise comparisons of the excess values determined by a payoff distribution x and the the excess values determined by a payoff distribution y: given two vectors x and y we need to be able to say that, say, x is greater or equal to y in some sense.
For a given payoff distribution there are as many excess values as there are coalitions, so we need to compare vectors in $\mathbb{R}^m$ with $m = 2^N$. As a consequence, we need to introduce at least a partial order on the space $\mathbb{R}^m$.
The formal definition relies on the lexicographic ordering of excess vectors from $\mathbb{R}^m$ and defines nucleolus and the the least vector, i.e., the one corresponding to the most stable allocation.
We show below an example of comparison and then the formal definition of nucleolus-

Example 7 (Excess comparison) *Suppose that TU game with characteristic function μ be played by the players $1,2$ and 3 and consider the payoff distribution vectors x and y*

- *given μ, create an m-tuple as follows* $(\mu(1), \mu(2), \mu(3), \mu(12), \mu(13), \mu(23), \mu(123))$
- *take the payoff distribution x and compute* $(x(1), x(2), x(3), x(12), x(13), x(23), x(123))$
- *take the payoff distribution y and compute* $(y(1), y(2), y(3), y(12), y(13), y(23), y(123))$
- *compute the excess vector for x, i.e.,* $e(x) = (\dots, (\mu(C) – x(C)), \dots)$
- *compute the excess vector for y, i.e.,* $e(y) = (\dots, (\mu(C) – y(C)), \dots)$
- *permutate the components of the vector $e(x)$ so as to order the components in decreasing order, call the new vector $e(x)^{\triangleright}$*
- *permutate the components of the vector $e(y)$ so as to order the components in decreasing order, call the new vector $e(y)^{\triangleright}$*
- *if either $e(x)^{\triangleright} = e(y)^{\triangleright}$, or in lexicographical order $e(x)^{\triangleright}$ comes before $e(y)^{\triangleright}$, we say that $x \preceq y$ (x precedes or equals y lexicographically), i.e., in our sense*

The apex $\triangleright$ applied to a vector denote the reordering of the vector in decreasing order. As a concrete example (from Stephane Airau [2]), consider the excess vector $e(y) = (-3, -3, -2, -1, 1, 1, 0)$ and the excess vector $e(x) = (0, 1, 0, 0, -2, -3, -3)$. We sort their components in decreasing order obtaining $e(y)^{\triangleright} = (1, 1, 0, -1, -2, -3, -3)$ and $e(x)^{\triangleright} = (1, 0, 0, 0, -2, -3, -3)$: the first components of the two vectors are equal, i.e., $e(y)_1 = e(x)_1$, but already the second components are different: $e(y)_2 > e(x)_2$. In this case $x \preceq y$, i.e., x is preferable to y.
The first entry of the sorted excess vector is, of course, the maximum excess: the players involved in the corresponding coalition have the largest incentive to leave their current coalition, and so on.
In this sense the nucleolus represents, among the possible configurations, the one which features the ”least problematic” set of complaints.

Definition 36 (Nucleolus) *Let Imp be the set of all the imputations. The nucleolus $Nu(\mathscr{N},\mu)$ is the set*

$$Nu(\mathscr{N},\mu) = \{x \in Imp \mid \forall y \in Imp \;\; e(y)^{\triangleright} \succeq e(x)^{\triangleright}\}$$

Summarising, the nucleolus relaxes the stability requirements of the core: the core admits no complaints, while the nucleolus may allow for some complaints, but tries to keep them at a minimum.
It is possible to prove that the nucleolus is always included in the core. This means also that if the core of a game in non-empty, one can use the nucleolus to discriminate between different core allocations.

We state without demonstration a result that guarantees that the nucleolus is non-empty in most games: indeed one can show (see for instance [11]) that if $Imp \neq \emptyset$ the nucleolus is non-empty. Furthermore another fundamental theorem, again given here without demonstration, states that the nucleolus consists in a unique payoff distribution.

Theorem 1 (Unicity of the nucleolus) *The nucleolus has at most one element.*

So in plausible games, the nucleolus is non-empty and unique. A major problem in exploiting the notion of nucleolus in computational game theory is that its computation is typically exponential in the number of players [14].

5.2.5 Definitions related to the Shapley Value

Actors in our coalitional setting are not only concerned with their individual economic advantage; rather they are typically sensitive of comparative distributional justice. In other words they value fairness, and their behaviour is negatively affected when they experience unfair treatment. Lack of fairness, as already discussed in the previous deliverable, can be considered as a cause for dysfunctional behaviour of an actor, in the present deliverable we will use perceived unfairness towards a group of actors to study also the possibility of dysfunctional behaviour of coalitions. Thus we recall the main concepts and introduce the necessary extensions.

The definitions of fairness referring to revenue distributions are several. A specific set of requirements broadly accepted as a definition of fair redistribution, and taken as axioms has been shown by Shapley [28] to correspond, given a game, to exactly one allocation: this allocation is called Shapley Value. Hereafter we recall the axioms and the Shapley solution.

We point, in passing, to the fact that in general the allocation x corresponding to the Shapley Value is not in the core. It belongs to the core for particular classes of games, such as convex games.

5.2.5.1 Fairness Axioms

Efficiency, Symmetry, Dummy player. Given a super-additive game with characteristic function/measure μ, we look for an allocation of values v_i such that

- **Efficiency** – It must hold
$$\sum_{i=1}^{n} v_i = \mu(N)$$
- **Symmetry** - If the two players i and j are *interchangeable*, i.e., if $\mu(S \cup i) = \mu(S \cup j)$ for every set not containing neither i nor j, it holds
$$v_i = v_j$$

- **Dummy player** – i is a dummy player if the amount that he contributes to any coalition is exactly the amount that i is able to achieve alone: if i is a dummy player he has to get a payment equal to exactly the amount that he would achieve on its own
$$v_i = \mu(i)$$
In other words there is no synergy between the player i and the other players.

Linearity Consider two different coalitional games, defined by two different characteristic functions μ_1 and μ_2, involving the same set $\mathcal{N}$ of agents.

- **Additivity** – If we re-model the setting as a single game in which each coalition S achieves a payoff of $\mu_1(S) + \mu_2(S)$, the agents' payments in each coalition should be the sum of the payments they would have achieved for that coalition under the two separate games.
$$v_i(N, \mu_1 + \mu_2) = v_i(N, \mu_1) + v_i(N, \mu_2)$$
where the game $(N, \mu_1 + \mu_2)$ is defined by
$$(\mu_1 + \mu_2)(S) = (\mu_1)(S) + (\mu_2)(S)$$
for every coalition S

5.2.5.2 The Shapley Value

Theorem 2 (Shapley Theorem) *There is a unique efficient payoff division (of the full payoff of the grand coalition) that satisfies the Symmetry, Dummy player and Additivity axioms:*
$$v_i^{Shapley} = \frac{1}{N!} \sum_{\sigma} \Big(\ \mu(\{C \cup i\}) - \mu(C) \ \Big)$$
where the index σ runs over all the $N!$ permutations of the N elements of $\mathcal{N}$.

An alternative expression is the following.
$$v_i^{Shapley} \ = \ \sum_{C \subseteq N} \frac{c!\,(N-c-1)!}{N!} \Big(\ \mu(\{C \cup i\}) - \mu(C) \ \Big)$$

Where the C's are the coalitions of $\mathcal{N}$ (the subsets of $\mathcal{N}$, i.e., $C \in 2^{\mathcal{N}}$) and $c = |C|$.
It is important to introduce here a remark on the Additivity axiom. In the words of Airiau [2], the additivity axiom, or ADD, is harder to motivate in some cases. If the valuation function of a TU game is interpreted as an expected payoff, then ADD is desirable (as you want to be able to add the value of different states of the world). Also, if we consider cost-sharing games and that a TU game corresponds to sharing the cost of one service, then ADD is desirable as the cost for a joint-service should be the sum of the cost of the separate services. However, if we do not make any assumptions about the games $(\mathcal{N}, \mu_1)$ and $(\mathcal{N}, \mu_2)$, the axiom implies that there is no interaction between the two games. In addition, the game $(\mathcal{N}, \mu_1 + \mu_2)$ may induce a behavior that may be unrelated to the behavior induced by either $(\mathcal{N}, \mu_1)$ or $(\mathcal{N}, \mu_2)$, and in this case ADD can be questioned."

For this reason several other equivalent axioms have been considered. A parsimonious set of axioms is due to Myerson [22] and it is based on a concept called balancing of contributions. Myerson considers the N restrictions of the original game μ, each defined by a single player leaving the game: a definition of value applied to each of those games has the balancing property if every pair of players the amount that each player wins or loses if the other leaves the game is the same.
Let us denote by $\mu \setminus i$ the game $(\mathcal{N} \setminus i, \mu_{\setminus i})$, where $\mu_{\setminus i}$ is the restriction of μ to $\mathcal{N} \setminus i$.

Definition 37 (Balanced contribution axiom) *A value function v satisfies the balanced contribution axiom iff for all $(i, j) \in \mathscr{N}^2$*

$$v_i(\mu) - v_i(\mu \setminus j) \;=\; v_j(\mu) - v_j(\mu \setminus i)$$

The new characterization of the Shapley Value due to Myerson is the following.

Theorem 3 (Shapley Value Theorem 2) *The Shapley value is the unique value function that satisfies the balanced contribution axiom.*

5.2.5.3 The Weighted Shapley Value

One of the main axioms that characterize the Shapley value is one of symmetry. In the words of Kalay and Samet [18]: ”The underlying motivation for using this axiom is the assumption that except for the parameters of the games, the players are completely symmetric. However, in many applications, this assumption of symmetry seems unrealistic for the situation that is being modeled and the use of nonsymmetric generalizations of the Shapley value was proposed in such eases. [...] It may be, for example, that a greater effort is needed on the part of player one than on the part of player two in order for the project to succeed. Another example arises in situations where player one represents a large constituency with many individuals and player two’s constituency is composed of a small number of individuals. Other examples where lack of symmetry is present can easily be constructed for problems of cost allocations. Also, lack of symmetry may arise when different bargaining abilities for different players are modelled.”.

Here, we are interested in the case in which a single player represents a constituency of more than one player, and the presence of a market (external to the grand coalition that takes part to the business process) for the contributions of the players may endow some player with different bargaining strength (within the grand coalition).

Algebraic approach to the Weighted Shapley Value. A weighted Shapley Value was defined by Shapley itself in the seminal work [28]. Consider the following games:

- the unanimity game $\mu = u_S$ w.r.t. a coalition S where $u_S(C) = 1$ if $S \subseteq C$ and $u_S(C) = 0$ otherwise
- the (canonical) game $\mu = w_S$ w.r.t. a coalition S where $w_S(C) = 1$ if $S = C$ and $w_S(C) = 0$ otherwise

It is well known that they form a basis for the space of coalitional games: the games w_S are the standard canonical basis of linear algebra.
Any game μ can be expressed in the canonical basis trivially as

$$\mu(C) = \sum_{S \subset \mathscr{N}} \mu(S) w_S(C)$$

and in the unanimity game basis as

$$\mu(C) = \sum_{S \subset \mathscr{N}} m_\mu(S) u_S(C)$$

where $m_\mu(S)$ is the Harsanyi dividend of S, related to μ by the Möbius transform

$$m_\mu(S) = \sum_{C \subseteq S} (-1)^{|S|-|C|} \mu(C)$$

and represents the genuine contribution of a coalition S to $\mu(S)$ that is not amenable to the contributions from the subsets of S; consequently one has also

$$\mu(C) = \sum_{S \subseteq C} m_{\mu}(S)$$

Let λ be a weight vector over a grand coalition N, i.e., a vector of positive real numbers whose N components reflect the weights of particular players.
The weighted Shapley value of the unanimity game is

$$v_i^{WShapley}(\mathcal{N}, u_S, \lambda) = \frac{\lambda_i}{\lambda(S)} \; \forall i \in S \quad v_i^{Shapley}(\mathcal{N}, u_S, \lambda) = 0 \; \forall i \notin S$$

This means that for any game μ by linearity one has the following expression for the weighted Shapley value

$$v_i^{WShapley}(\mathcal{N}, \mu, \lambda) = \lambda_i \sum_{S: i \in S} \frac{m_{\mu}(S)}{\lambda(S)} \qquad \forall i \in \mathcal{N} \tag{5.1}$$

this expression is in dividend form.
The corresponding expression in coalitional form is

$$v_i^{WShapley}(\mathcal{N}, \mu, \lambda) = \lambda_i \sum_{S: i \in S} \gamma_S \Big[\; \mu(S) - \mu(S \setminus \{i\}) \; \Big] \qquad \forall i \in \mathcal{N} \tag{5.2}$$

where

$$\gamma_S = \sum_{C: C \cap S \neq \emptyset} \frac{(-1)^{|}C|}{\lambda(C \cup S)}$$

Obviously the original expression of the Shapley value is obtained for $\lambda_i = 1 \, \forall i$, and $\lambda(S) = |S|$.

Example 8 (Weighted Shapley value) *Consider a game in coalitional form*

$$\begin{aligned} \mu(1) &= 100 \\ \mu(2) &= 200 \\ \mu(3) &= 300 \\ \mu(1,2) &= 400 \\ \mu(1,3) &= 500 \\ \mu(2,3) &= 600 \\ \mu(1,2,3) &= 900 \end{aligned}$$

The (ordinary) Shapley Value is

$$v_i^{Shapley}(\mathcal{N}, \mu) = (200, 300, 400)$$

Let us take the weight vector

$$\lambda = (\frac{1}{8}, \frac{3}{8}, \frac{1}{2}).$$

To compute the Weighted Shapley value $v_i^{WShapley}(\mathcal{N}, \mu, \lambda)$ in coalitional form, compute first the γ's:

$$\gamma_1 = \frac{27}{5}, \qquad \gamma_2 = \frac{11}{21}, \qquad \gamma_3 = \frac{9}{35},$$

$$\gamma_{12} = 1, \qquad \gamma_{13} = \frac{1}{5}, \qquad \gamma_{23} = \frac{1}{7},$$
$$\gamma_{123} = 1$$

then we obtain

$$v_i^{WShapley}(\mathscr{N}, \mu, \lambda) = (145, \frac{2225}{7}, \frac{3060}{7})$$

Using the dividend form, we would obtain the same results by computing first the dividends

$$m_1 = 100,\; m_2 = 200,\; m_3 = 300, \qquad m_{12} = m_{13} = m_{23} = 100, \qquad m_{123} = 0\,.$$

Example 9 (Cardinality Weighted Shapley value) *A relevant example is the one in which the grand-coalition is partitioned into a number of coalitions, i.e., it is endowed by a coalitional structure $CS = \{S_1, S_2, \ldots, S_k, \ldots, S_m\}$ with $S_j \cap S_k = \emptyset\ \forall j,k$ and $\cup_k S_k = \mathscr{N}$. In some cases the relationship among members of a coalition S_k are such that the issue of fairness is relevant not only at individual player level, but also at coalition level (this is the case for instance of solidarity relationships). In this case one can compute the Shapley Value of each coalition using $\lambda_k = |S_k|$.*

Random-order approach to the Weighted Shapley Value. The approach used above is known as the algebraic approach to the Weighted Shapley Value: the value of a unanimity game is determined first, by allocating one unit among the players of the carrying coalition according to their relative weights, then the value for general games is obtained by linearity from the unanimity base. An alternative approach is the random order approach: there the weights are used to determine a probability distribution over orders of the players, then the Weighted Shapley Value is defined as the expected contributions of the players according to this probability distribution.

5.3 Applications of Coalition Theory to the Risk Management Framework

5.3.1 Definitions

We will now apply the notions of the previous sections to enrich our risk management framework. We recall that the risk associated to an adverse event E can be given a quantitative formalisation by taking into account two components: the damage (*impact*) incurred by the actors in the case of occurrence of the adverse event, the likelihood of occurrence of the adverse event. The impact to a participant A from the event E can be written $I(A,E)$.

The *likelihood* of the event takes most often the form of a probability $Pr(E)$. It is a quantity that in turn depends on the probability of other triggering events and conditions to take place. It can be computed based on the definition of the collaborative process, and/or quantified in the basis of expert opinion.

Overall the risk from an adverse event E to a party A is

$$R(A,E) = I(A,E) \times Pr(E)$$

The risk management process consists of two logically distinct phases: the *risk assessement* (risk identification, risk analysis, evaluation of relevant risks); the *risk treatment* (identification of options, action plan development and implementation, identification and analysis of residual risks).

5.3.2 Risk assessment

In our framework, the assessment phase is an iterative sub-process inserted within the standard risk management process (see the previous chapter). It includes: (i) surveying the possible disclosure events and estimating their impacts; (ii) analysing the process so as to identify how the combination of factors could trigger the disclosure event E; (iii) estimating the probability of the individual component factors of an event E (iv) aggregating them into an estimate of $Pr(E)$, and finally ranking the different risks.

5.3.2.1 Input information

We assume that the risk assessment sub-process receives in input the following information:

- A specification of the business process where risk needs to be assessed, consisting in
  - Purpose and Scope
  - Process roles and actors
  - Process input and expected output
  - Process representation (in the form of the syntax discussed in the previous chapter), including exceptions
  - Information items owned, exchanged or produced during the process
  - Availability of information items to the individual actors during the process steps

5.3.2.2 Adverse events survey and impact assessment

As discussed in the previous chapter, we consider adverse events that belong to the categories of information leakage. Our approach is readily applicable to other insider attacks (i.e.,, attacks carried out by participants in the business process), such as information misrepresentation and out-of-time communication threats (including missed communication). Specifically:

- for information leakage threats, we identify one or more actors to which an information item is available at some step of the process and who
  - could leak that information to third parties outside the business process, or
  - could pass the information to other participants in the process that are not intended to receive that information;
- for out-of-time communication attacks, the sub-process generates the scenarios where one or more information items are not communicated in due time (for instance in a collaborative forecasting process a vendor omits to communicate his recent sale volumes or the expected future sale volumes);
- for misrepresentation attacks, the sub-process generates representatives of those scenarios where the value of a quantity is altered before communication (for instance in a collaborative forecasting process a vendor inflates his recent sale volumes or the expected future sale volumes);

We remark that in principle in all the above types the attack can be performed by individual participants or by coalitions. However coalitional out-of-time communication attacks can be formally reduced to sets of individual out-of-time communication attacks. Thus coalitional attacks we consider here are

- coalitional information leakage.
- coalitional misrepresentation attacks

For instance in a supply chain, same type actors not only contribute to the overall chain, but compete also with one another: in that case coalitional misrepresentation attacks correspond to agreements among sets of actors to misrepresent their aggregated attributes and distort competition within the chain.

Of special interest is coalitional information leakage. Some times individual leakage is not viable, because information in exchanged messages is protected, but coalitions can work out the information: this is, for instance, the case of information protected by a k out of n secret sharing schema, where individual actors have the availability of a single share. The leakage that cannot be performed by individual actors can be performed by a coalition of at least k actors.

The survey of (individually or coalitionally generated) adverse events, is followed by an estimate of their impact to the actors $I(A,E)$. Again this part of the analysis can be supported by our game-theoretical models: for instance, the loss of competitiveness by actors within a Supply Chain following from actors' collusion can be estimated also by quantitative models of the specific SC system.

Once the risk assessor has filled out the catalog of the relevant adverse events E along with their impact on each actor and with the indication of the originating actor or coalition, she can proceed to the analysis of the likelihood of attack, by individuals and by coalitions respectively.

5.3.2.3 Individual actor risk assessment

The dysfunctional behaviour of the individual actor can be caused by unsatisfaction due to perceived unfairness or by greed (e.g. due to perspective illegal gains obtainable by illicitly selling the information). We adopt here a probabilistic approach: the likelihood of misbehaviour is understood in terms of frequency-based probability.

The perceived unfairness driver. Perceived unfairness can be quantified by computing for each player i an index ϕ_i of unfairness in the process resource allocation, then associating to each player a probability of dysfunctional behavior. Now let $v_i^{factual}$ be the payoff to an actor. If the difference between this quantity and the actor's Shapley Value (as defined in the previous Section)is positive, the actor is under-rewarded for her contribution: this situation may feed its propensity toward a defection; if, instead, this difference is negative, the actor is over-rewarded and the discrepancy does not contribute to its propensity towards defection. One needs also to relate the discrepancy, to the absolute value of $v_i^{Shapley}$. For all the above considerations, the dissatisfaction parameter ϕ_i for an actor i can be defined as follows:

$$\phi_i \equiv \theta\left(v_i^{Shapley} - v_i^{factual}\right)/v_i^{Shapley}, \tag{5.3}$$

where $\theta(z)$ is such that $\theta(z) \equiv 0$ if $z < 0$ and as $\theta(z) \equiv z$ otherwise. The index ϕ belongs by definition to the real interval $[0,1]$. Starting from ϕ_i one can obtain a probability of dysfunctional behavior $p_i = p(\phi_i)$ by a mapping provided by expert opinion. The function $p(\phi)$, whose precise analytic form can be obtained from the experts, must be a non-decreasing function starting from $p = 0$ at $\phi = 0$ and equal to $p = 1$ at $\phi = 1$ (which implies $v^{factual} = 0$).

The greed driver. A player receives as a consequence of the participation in the business process a prescribed amount. However the information she has access to might have some value for actors who are not intended to access that information. Those actors may belong to the business process or be external to the business process, e.g. they may be competitors. By selling that information the player can achieve an illicit profit. If not detected this action may bring to the player an economical advantage. Greed is a strong driver for human behaviour in a business process. The players' greed can be modeled based on the expected gain from an attack, which, as in a bet, is defined as the difference between the positive expected payoff obtained from a successful attack (for instance consisting in selling on an illegal market the information acquired) and the negative expected payoff associated to an exposure of the attack (job loss, reputation loss, fines or other punishments): this difference yields a balance, possibly incorporating risk aversion or risk seeking attitudes and correspond to a perceived utility b, that is then used to compute for each actor a probability $q_i = q(b_i)$ of defection. Also this function is an non-decreasing function of the argument. Let us now consider the example of a participant A to a business process who gets a peek to some information she's not supposed to see and considers selling this information to a third party outside the business process itself. More specifically, let us assume A has access to an information item I, whose value on the illegal market is $V(I)$; suppose that the probability that A can sell this information when she tries to do so is one. Suppose that if the illegal deal is exposed, A will incur in a punishment D and that this happens with probability L. Let be $u(\cdot)$ the utility function: if the actor is risk-neutral $u(v) = v$, if A is risk-averse $u(v)$ is concave, if A is risk-loving $u(v)$ is convex. Then the perceived expected value of the gamble from the point of view of A is

$$b = u(V(I)) - L \times u(D) .$$

If this balance is known it can be transformed into a probability by means of a mapping $q(b)$. The elements involved into this computation and their sources are the following:

- $V(I)$, the value of the information sold on the illegal market is provided by domain experts;
- L, the probability of exposure, can be either derived from the definition of the process (can the ex-filtration be tracked?) or by domain expert knowledge;
- D, the damage to the perpetrator, can be obtained by the regulations of the process
- $u(\cdot)$, the actor utility function, can be obtained by domain experts: if no specific information about an actor is known in this respect, one can assume risk-neutrality;
- $q(b)$, the expected utility to probability mapping, is provided by domain experts: it starts from $q = 0$ at $u = 0$ and it is an increasing function.

Composition of the drivers for the individual player. When a player A can be considered as sensitive to both drivers, perceived unfairness and greed, the probability of attack P_i is worked out from the following "series-system" relation

$$(1 - P_i) = (1 - p_i)(1 - q_i)$$

which expresses the fact that the reliability of an actor is the product of her reliability w.r.t. the different factors. In this expression the actions of the two drivers is considered, for sake of simplicity, uncorrelated.

5.3.2.4 Coalition risk assessment

In a business process, groups of actors, i.e., subsets of $\mathscr{N}$, might form coalitions different from the grand coalition $\mathscr{N}$ (sub-coalitions). Sub-coalition behaviour may deviate from the globally agreed protocol, with the aim of increasing the benefits to the sub-coalition (to be later shared among the sub-coalition members). The analysis of the likelihood of dysfunctional behaviour by groups of actors is thus, in some respect, conceptually analogous to the one of individual actors. In the case of individual actors we considered two main drivers, i.e., perceived unfairness (in the redistribution of the returns of the grand coalition) and greed. The same two drivers, in the interpretation clarified below, can be introduced for coalitions.

- Greed can be a driver for a coalition in that a collaborative dysfunctional activity can bring an undue benefit (not obtainable by individuals alone) that later is shared among coalition members.

- Perceived unfairness too can be a driver of a coalition, when the corresponding set of actors is tied together by bonds such as professional solidarity, gender solidarity, same-language solidarity, same-region solidarity and the like. When this kind of bonds are in place, the grand coalition can partition into smaller coalitions, i.e., assume a coalitional structure.

We consider the perceived unfairness driver first (an approach based on the Weighted Shapley Value), then the greed driver (we define an approach based on the core concept and the gamble paradigm used for individual actors).

Unfairness towards a coalition When a coalitional structure is present, each coalition making up the partition can be considered an individual player, with a weight corresponding to the Choquet capacity of the coalition. The approach used to model this issue develops along the lines of the Shapley Value based method, used for the individual actors. However, in place of the ordinary Shapley Value, we use the Weighted Shapley Value, defined in equation (5.2): the likelihood of defection is computed based on the discrepancy of the total factual payoff of the coalition members and the weighted Shapley Value of the coalition computed by using the coalition cardinality as a weight. Perceived unfairness at coalition level is quantified by computing for each coalition C an index ϕ_C of unfairness in the process resource allocation, then associating to each coalition a probability of dysfunctional behaviour due to this driver.
Assume a non-trivial coalitional structure CS is given (e.g. $CS = \{C, (\mathscr{N} \setminus C)\}$). Now let $v_C^{factual} = \sum_{i \in C} v_i^{factual}$ be the total payoff to the coalition, furthermore let $v_C^{Shapley}$ be the Shapley Value of the coalition computed according to equation (5.2). If the difference between this quantity and the coalition's Shapley Value is positive, the coalition is under-rewarded for its contribution.
One can observe, incidentally, that the perceived unfairness driver for a coalition is related to the the perceived unfairness driver for its members: on the one hand, obviously, if all the members are under-rewarded, also the coalition is under-rewarded; on the other if the total factual payoff of the coalition members is lower than the Shapley value of the coalition, at least a member is under-rewarded. The dissatisfaction parameter $\phi_C \in [0,1]$ for a coalition C is defined as:

$$\phi_C \equiv \theta\left(v_C^{Shapley} - v_C^{factual}\right)/v_C^{Shapley}, \tag{5.4}$$

where $\theta(z)$ is the Heaviside step function ($\theta(z) \equiv 0$ if $z < 0$ and as $\theta(z) \equiv z$ otherwise). Starting from ϕ_C one obtains a probability of dysfunctional behavior $p_C = p^{coal}(\phi_C)$ by a mapping provided by expert opinion. The analytic form of the function $p^{coal}(\phi)$ can be obtained from domain experts (also in this case, it must be a non-decreasing function starting from $p^{coal} = 0$ at $\phi = 0$ and equal to $p^{coal} = 1$ at $\phi = 1$).

Greed driver. A coalition, like an individual, can be driven by greed and adopt a dysfunctional behaviour, if this can result in an expected payoff that exceeds the one pre-agreed with the other members of the grand coalition. Indeed, the analysis of the greed driver for a coalition is analogous to the one performed for individuals: the former generalises the latter by using group in place of individual rationality.

We are now ready to apply the notion of Core introduced in the previous section. For the grand coalition, the core is the set of efficient redistributions of payoffs that are both individually rational and group rational: no individual nor group would be better off by splitting. Our method considers allocations lying outside the core to be critical: the corresponding individuals and coalitions are likely to act dysfunctionally.

A key element, here, is that coalition level greed is mostly aimed at satisfying individual greed of the coalition participants: the illicit outcome should eventually be shared among members. The very fact of outcome sharing can be used for contrasting this driver. On the one hand, as individual greed can be contrasted by reducing the expected value of the gamble from the point of view of the individual, coalition level greed can be contrasted by reducing the expected value of the gamble from the point of view of the coalition; on the other hand, coalition greed can be contrasted by acting on the expectations of individual components. For instance the threat of attack from a specific coalition can be dealt with, by conveying incentives toward target members, so that they are not interested in pursuing their participation into the illicit group action (this represents a sort of ”*divide et impera* approach).

5.3.3 Risk treatment

According to the ENISA [1] ”Risk Treatment is the process of selecting and implementing of measures to modify risk. Risk treatment measures can include avoiding, optimizing, transferring or retaining risk.” We point to the general document [1], for details. In the present context, where the adverse events are information leakage attacks or information mis-representation attacks, the security controls to be used as risk treatment options depend on the specific use case at hand. The PRACTICE project has defined a rich panoply of such controls. We make some examples

- Risk avoidance can be achieved by cryptographic controls: for instance the risk of unintended release of sensitive information used in a computation can be avoided by enacting the computation within a homomorphic encryption scheme.

- Risk mitigation can be achieved by reducing either the impact or the probability of the adverse events, for instance the probability of events triggered by humans that are motivated by lack of satisfaction can be reduced by the use of incentives.

- Risk transfer can be achieved by means of insurance policies or other financial instruments (this case is not considered further in this deliverable)

- Risk retention can be acceptable when the cost of the countermeasures exceeds the available resources: it may be the case that for instance no further reduction of the probability of adverse event is possible. Typically the choice of which risks to retain is based on the ranking of the relevant risks.

5.4 Use case: a vendor managed inventory system

We now consider a use case consisting in the forecast of the aggregated demand, in two reference settings, defined by different responsibility allocations among parties: the Customer Managed Inventory (CMI) system (also referred to as Normal Replenishment System) and the Vendor Managed Inventory (VMI) system. This use case is an improved version of the one presented in Deliverable 31.2, including the coalition analysis notions introduced in the previous section of the current deliverable.

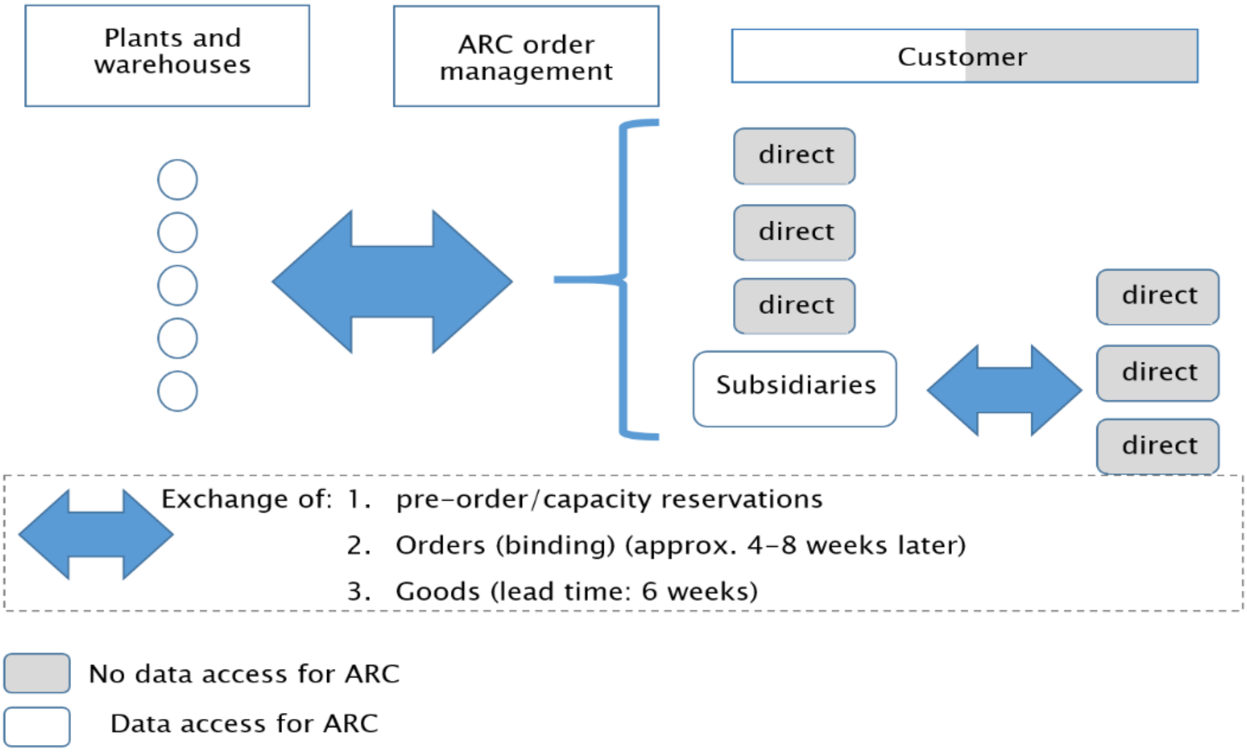


Figure 5.1: A schema illustrating the ARC supply chain: in this case the Vendor is ARC.

The players in both reference settings are the Vendor and its Customers. Customers are typically large retail stores that work independently, the Vendor is their exclusive wholesaler.
In the CMI setting, Customers manage their own inventory by forecasting their own local demand and then send their forecast to the Vendor; in the VMI setting the Vendor, takes the responsibility for managing the inventory of its Customers (see Deliverable D24.2 page115, see also Figure 5.1)). We briefly recall the terms of the problem. In both settings, the Vendor obtains the merchandise from a set of producers within its organisation: making the merchandise, ordered by the Customers, available to them, takes several weeks. From this the need arises of anticipating the aggregated demand by an estimate, provided by some forecasting process.

The AGGPRE process. Currently, the CMI setting is adopted and a simple demand-forecasting process is in place. This process, hereafter for the sake of conciseness called AGGPRE (aggregation of pre-orders), consists of two steps:

- the Customers send to the Vendor non-binding pre-orders 4-8 weeks before the binding orders; (in principle the pre-orders are meant to represent faithfully the individual customers demand forecasts);
- the Vendor computes the total of the pre-orders and adopts it as forecast value for the aggregated demand.

This method turns out to be imprecise and inaccurate, i.e., there is most often a significant discrepancy between the pre-orders and the actual binding orders.
The prediction error is not only an effect of the market fluctuations – and of the consequent imprecision of the individual Customer forecast – but mostly depends on the selfish behaviour of the Cus-

tomers: each customer has incentives to inflate pre-orders, in order to prevent shortage of merchandise in the event of a market increase: this introduces a positive bias and compromises the accuracy of the forecast. The resulting forecast distortion can be considered a misrepresentation attack delivered by the Customers.

The COLFOR-LR process. In Deliverable D24.2 the VMI setting is proposed, supported by a more accurate forecasting mechanism: a collaborative forecast using linear regression – from now on called COLFOR-LR (collaborative forecast by linear regression):

- each Customers sends to the Vendor her own historical record of the sales (the Point of Sale or POS data): notice that normally this is private information to the Costumer
- based on those data, the Vendor performs the forecast of the aggregated demand using linear regression.

The COLFOR-LR forecasted values are not affected by the customers selfishness, therefore they can, in principle, be more accurate than those from AGGPRE. Notice that within COLFOR-LR, pre-orders could still be issued by customers as complementary information.
An obvious concern in COLFOR-LR is keeping the historical POS information private to the owners, i.e., the Customers. In fact, that information, if openly made available to the Vendor could be used to distort the internal supply chain market. The historical information of a specific Customer

- would offer to the Vendor supplementary bargaining power and,
- if passed to competitor Customers, it could be used to distort competition.

The COLFOR-LRSEC process. Deliverables D24.2 and D24.3 address this issue: aggregated demand forecasts that keep the input information private can be derived by performing secure linear regression analysis of the POS based on a secure collaborative forecasting and planning system deployed on a cloud computing facility (see Figures 5.2 and 5.3 from Deliverable D24.2, Section 3.2.2 Collaborative forecasting for mid-term production planning). This secured process, hereafter called COLFOR-LRSEC (for Secure Linear Regression) is enacted by three kind of actors: one Vendor, at least one Cloud actor, and several Customers. The process operates as follows:

- each Customer sends to the Cloud Actor her own historical POS record of sales;
- based on those data, the Cloud Actor performs the forecast of individual Customer's sales volumes using linear regression and aggregates the demand; the operations are performed using homomorphic encryption;
- then the Cloud Actor sends the aggregated demand to the Vendor.

In the process, none of the parties can learn the input data of another individual party neither can the Cloud actor.

However, another issue arises in the VMI processes (both by COLFOR-LR and COLFOR-LRSEC): the reduction of the prediction error obtained collaboratively in the aggregated demand forecast goes mainly to benefit of the Vendor:

- on the one hand the Vendor experiences a risk reduction,

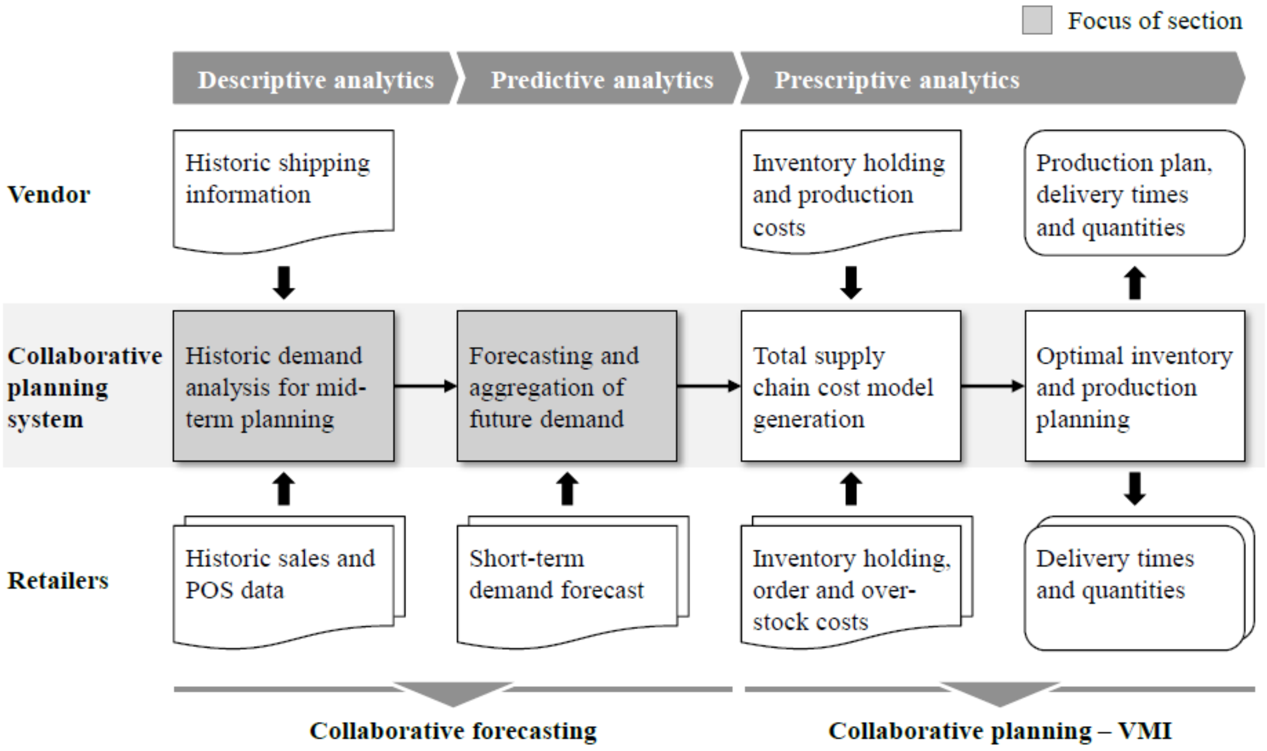

Focus of section
Descriptive analytics
Predictive analytics
Prescriptive analytics
Vendor
Historic shipping information
Inventory holding and production costs
Production plan, delivery times and quantities
Collaborative planning system
Historic demand analysis for mid-term planning
Forecasting and aggregation of future demand
Total supply chain cost model generation
Optimal inventory and production planning
Retailers
Historic sales and POS data
Short-term demand forecast
Inventory holding, order and over-stock costs
Delivery times and quantities
Collaborative forecasting
Collaborative planning – VMI

Figure 5.2: Processes on focus.

- on the other hand the Customers, no longer able to inflate their pre-orders, risk merchandise shortages, coming from a conservative forecast.

The AGGPRE process and the COLFOR-* process are characterized by opposite unbalancing of the risk coming from the two – opposite signs – events, of overestimation and underestimation of market demand:

- in the AGGPRE process, within a Client Managed Inventory setting, estimates are tendentially optimistic, i.e., they tend to assume a higher market; therefore the most likely adverse event is the overestimation of the aggregated demand; this event affects mostly the Vendor, which is left with unordered merchandise

- in the COLFOR-* process, within a Vendor Managed Inventory setting, estimates are tendentially conservative; in typical commercial arrangements an underestimate affects mostly the Customers

This can be considered unfair by the Customers. As a consequence, they could

- refuse altogether to participate to COLFOR-*,

- or they could agree to participate to COLFOR-*, but tamper with their own data, before those are fed in input to the forecasting system; in other words, Customers, can still deliver a misrepresentation attack: instead of distorting the individual forecasts, they distort the input data of the collaborative forecast.

We will see how the risk of this dysfunctional behaviour by the customers can be captured by the Shapley Value based analysis introduced above.

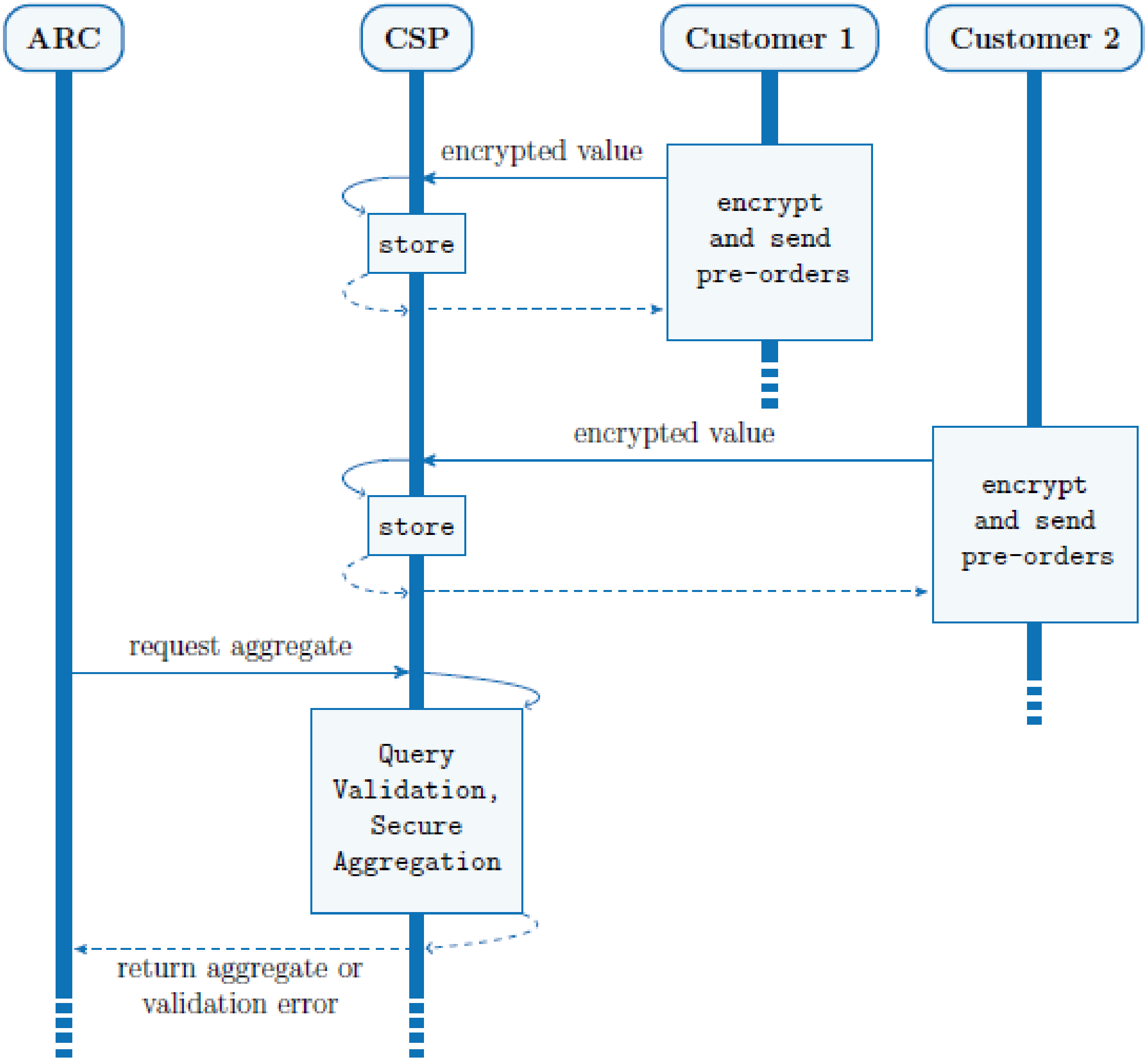

ARC
CSP
Customer 1
Customer 2
encrypted value
store
encrypt and send pre-orders
encrypted value
store
encrypt and send pre-orders
request aggregate
Query Validation, Secure Aggregation
return aggregate or validation error

Figure 5.3: Sequence diagram and data flow of the COLFOR-LRSEC forecasting process. ARC is the Vendor, CSP is the Cloud Actor, Customer1, Customer2, ect. are the Vendor's Customers.

However, already from the problem set-up of we can see that countermeasures to this undesired behaviour might consist of incentives, for instance a market-volatility-risk revenue redistribution among Vendor and Customers.

5.4.1 Risk assesement

We consider the case of the demand forecast of a product, in the context of the Arcelik Supply Chain which has been analyzed in WP24 and focus the analysis on a particular item produced, a washing machine, whose reference number is denoted with $SKU7128644900$: simulations based on real data (see Deliverable 24.5), show that collaborative forecasting using linear regression can reduce the prediction error.

Specifically, passing from AGGPRE to COLFOR-* with faithful input, one can reduce the error on the result by about 10%; that this improvement, in turn, translates into a cost reduction for the Supply Chain of 12%.

Recalling once more the business process risk definition as product of impact and probability $R(A,E) = I(A,E)Pr(E)$ (where E represents the adverse event and A a participant to the process), we estimate first the probability and the impact of the dysfunctional behaviour consisting in misrepresentation attacks, where Customers tamper with the values of their historical data so as to inflate the individual Customer forecast.

5.4.1.1 Impact

It is reasonable to assume that, overall, the Customers considering to tweak the COLFOR-* forecasting process aim at achieving the same inflated levels they would use in AGGPRE. We assume therefore that the misrepresentation attack can produce a degradation of the accuracy of the order of 10% and into a cost increase for the Supply Chain of 12%, mostly incurred by the Vendor, in a typical VMI setting. Thus, indicating the event of misrepresentation attack by the Customers by E, we have

$$I(Vendor, E) = c \times K$$

where K is the cost to the Vendor for the provision of the marchandise and r is the percentage increase, in our case estimated to be approximately $r \simeq 0.12$.

5.4.1.2 Probability

The probability of the misrepresentation attack can be related in this case to the perception of unfairness from the Customers. Note that this type of attack does not involve any risk from the side of the Customer, since, in the event of a market overestimate issued by COLFOR-*, it is difficult to prove the private data tampering by the individual Customers: the attack will remain an hidden action.
To quantify the probability of this attack we perform the Shapley analysis at individual level. The reason why, in this case, we do not perform a Shapley analysis at the coalitional level is that the attack we are considering is delivered by each individual Customer with no need for involving other Customers: each individual Customer aims at an individual objective and the overall attack consists in the simple sum of the attacks from individual Customers.
In this case study we also assume that the unfairness related analysis does not apply, to Cloud Actors because, typically, they are contracted by the Vendor to provide the computation service and are not part of the Supply Chain collaboration. Furthermore, since Cloud Actors are paid for a service that is independent of the accuracy and precision of the forecast, they do not incur in any benefit or any risk

arising from the forecast error. Fairness in the present process concerns only Vendor and Clients. Consider, therefore, the collaborative process aimed at reducing the error of the demand forecast. The players of this process are of two kinds: Vendors and Customers; there is a single Vendor, denoted by V, and there are m Customers, denoted by $C_1, C_2, \ldots, C_m$, that make up the set $\mathscr{C} = \{C_i\}_{i=1}^{m}$. The overall player set is $\mathscr{N} = \{V, \mathscr{C}\}$ and overall there are $N = m + 1$ players.
We also denote by $u_1, u_2, \ldots, u_m$ the respective demand volumes of the Customers: those volumes can be obtained by forecasting the individual demand of each Customer, based on correct data on the Customer history.
The Vendor is central to the forecasting process in that she triggers the forecasting algorithm and provides historical shipment data; in our setting she also pays for the Cloud service.

5.4.1.3 Quantification of the Shapley Values

To compute the Shapley Value of each player we need to quantify the contributions of the different sets of players to the final result (the reduction of the error prediction), i.e., we need to define the measure μ on the power set of all actors (find the worth of any set in $2^{\mathscr{N}}$). We observe in passing that when the actors play correctly, μ is a monotonically increasing function: the addition of a player to a set of players either allows to trigger the process (when the added player is the Vendor) or increases the information available for the forecast (when the added player is a Customer).

General considerations. In order to quantify the contribution of individual actors we observe the following.

- The Vendor alone can trigger the forecast but cannot reduce the error of the forecast, thus $\mu(V) = 0$
- Any set of players not containing the Vendor, i.e., containing only Customers, cannot trigger the forecast procedure, therefore in those cases the worth of the coalition C is zero

$$\mu(C) = 0 \qquad \forall C \in 2^{\mathscr{C}}$$

- Assuming a Vendor is present in the coalition, the influence of individual Customers on the reduction of the prediction error varies from one customer to another, however, in our setting, it is reasonable to assume that at the first order of approximation it is proportional to the demand volume of the Customer and that is additive. These assumptionas can be justified as follows.
  - A Customer has two choices: contribute his historical data or not; when the data are not contributed, the Customer formulates a guess about the future orders, i.e., issues a pre-order.
  - We know that pre-orders tend to be biased in excess (i.e., inflated): due to its nature this bias is a systematic error, thus by contributing historical data in place of issuing pre-orders (that is participating to the forecast process), Customers help reducing the systematic error, by reducing the excess.
  - Having always the same direction this correction is additive.
  - Specifically, the forecasted aggregated demand volume U for a product, is computed simply as the sum of the forecasted demand volumes u_i for individual customers: since –

conditional to the market – the demand volumes of the individual Customers can be considered independent from one another, one has

$$U = \sum_{i=1}^{m} u_i$$

If none of the information from the Customers is available the value of the aggregated forecast is set to a value $U*$ based on the pre-orders u^* sent by the Customers (which – we assume – encompass, implicitly, the background information held by the Vendor, such as the historical shipment records)

$$U^* = \sum_{i=1}^{m} u_i^*$$

- the difference $\Delta U \equiv U^* - U$ is the total error of the forecast, whereas the difference $\Delta u_i \equiv u_i^* - u_i$ is the individual error of the forecast

- We know from simulations the order of magnitude of the contribution and assume

  $$r = \Delta U / U \simeq 10\% \qquad \text{and} \qquad \Delta u_i / u_i \; \forall i$$

  every customer brings a contribution of the order of $r = 10\%$ of his demand volume in the reduction of the error.

- Being the bias always an excess, this contribution is additive.

As a final observation, we point to the fact that the demand volume of the individual Customers can be roughly estimated, lacking other sources, on the basis on the shipment logs, available to the Vendor.

The set measure μ of this game. In short, in the coalition $\{V, \mathscr{C}\}$ (where V is the Vendor and $\mathscr{C}$ contains alla the Customers), the added value of the actor i is

$$\Delta_i(\mathscr{N}) = \mu(\mathscr{N}) - \mu(\mathscr{N} \setminus i) = \Delta u_i = (U - U^*)\frac{u_i}{U}$$

On the other hand, we can adopt the same reasoning for all the coalitions containing V, therefore, exploiting additivity:

$$\mu(V \cup C) = r \sum_{i \in C} u_i \qquad \text{where } C \subseteq \mathscr{C}$$

while, when V is not present

$$\mu(C) = 0 \qquad \text{where } C \subseteq \mathscr{C}$$

This attributes a value $\mu(C)$ to every possible coalition C.
Thanks to this, one can compute the Shapley value $v_i^{Shapley}$ for each player.
The computation is considerably simplified by the additivity of the measure μ conditional to the presence of the Vendor.

The Shapley Value in an illustrative case. Let us illustrate the computation using an illustrative player set $\mathscr{N} \equiv \{ V = A, C_1 = B, C_2 = C \}$ consisting of the Vendor – we call A – and two Customers – that we call B and C respectively. Let us indicate, for the sake of notational simplicity, the demand volume B and C respectively by b and c. The usual permutation and added values table is the following

	A	B	C	
ABC	0	b	c	
ACB	0	b	c	
BAC	b	0	c	
CAB	c	b	0	
BCA	$b+c$	0	0	
CBA	$b+c$	0	0	
				Grand total
Total	$3(b+c)$	$3b$	$3c$	$6(b+c)$
Shapley Value	$\frac{1}{2}$	$\frac{1}{2}\frac{b}{b+c}$	$\frac{1}{2}\frac{c}{b+c}$	

One can observe that when A is in the first position his added value is 0, when it is second he gets a value equal to the volume of the first player (a Customer), when it is third he gets a value equal to the volume of the first two players (the two Customers).

The case with m Customers. Generalizing it is easy to see that the vendor can be in any position of the permutation with the same probability, thus in average over the permutation of the reminder elements, he gets a value equal to half of the demand: his Shapley value is always $1/2$ irrespective of the number m of Customers. The other players are left with half of the total value created by the collaboration, to be shared in parts proportional to the individual demands. Thus given the set of players $\{V, C_1, c_2, \dots, c_m\}$ and given the estimates u_i of their demands the Shapley Values are

$$v_V^{Shapley} = \frac{1}{2} \tag{5.5}$$

$$v_{C_i}^{Shapley} = \frac{1}{2}\frac{u_i}{U} \tag{5.6}$$

5.4.1.4 From the Shapley Value to the probability of misrepresentation attack

Even without the precise values, the positivity of everyone's Shapley Value is enough to notice the unfairness of the revenue reallocation in the present case. We have seen that the revenue of this process translates into savings by the Vendor: the other players do not benefit from the process. Let us indicate the revenue by $R = \Delta U$. The factual revenue distribution for the players $(0, 1, 2, \dots, m)$, where player 0 is the Vendor, is

$$v^{factual} = (v_0^{factual}, v_1^{factual}, v_2^{factual}, \dots, v_m^{factual}) = (R, 0, 0, \dots, 0)$$

the ideal, i.e., Shapley value, revenue distribution is

$$v^{Shapley} = (v_0^{Shapley}, v_1^{Shapley}, v_2^{Shapley}, \dots, v_m^{Shapley})$$

with $0 < v_0^{Shapley} < R$ and $v_1^{Shapley} > 0$ for all i's; more specifically the Shapley Value is indicated in equations (5.5)–(5.6).
Now the index of perceived unfairness can be computed.

We recall equation (5.3): the parameter ϕ_i for an actor i is be defined as:

$$\phi_i \equiv \theta\left(v_i^{Shapley} - v_i^{factual}\right)/v_i^{Shapley}, \tag{5.7}$$

here $\theta(z)$ is such that $\theta(z) \equiv 0$ if $z < 0$ and as $\theta(z) \equiv z$ otherwise. Thus $\phi_0 = 0$, i.e., the Vendor does not perceive the redistribution as unfair, while it is easy to see all that the other actors do: $\phi_i = 1$.

We have said that starting from ϕ_i one gets the probability of dysfunctional behavior $p_i = p(\phi_i)$ by a mapping whose precise analytic form can be obtained from the experts, but that must be a non-decreasing function starting from $p = 0$ at $\phi = 0$ and equal to $p = 1$ at $\phi = 1$.

In our case $p_0 = 0$ and $p_i = 1 \, \forall i > 0$: player V does not attack, while all the other attack with certainty, i.e., in this case, all the Customers will misrepresent the input data to some degree.

5.4.1.5 Risk computation and ranking

The estimate of the overall impact has been already addressed in Subsection 5.4.1.1. At this point we can compute the risk: the probability of an attack being equal to 1 for all the actors the overall risk is equal to the impact.
Since, the overall adverse event consists of the independent attacks one can establish a ranking of the components of this risk. It is reasonable to assume that the impact of each attack will be proportional to the demand volume of the Client. If the misrepresentation attack is certain, also the risk will be proportional to such a volume.

This volume can be estimated based on the shipment history, available to the Vendor: a precise value for this estimate is not necessarily needed when dealing with risk ranking.
The ranking of Customers by historical shipment volumes provides a suitable ranking of the risk associated with individual Customers.

5.4.2 Risk treatment

For addressing risks there are different options: avoiding, alleviating, transferring or retaining risk. In the present case the certainty of attack by Customers in standard conditions suggests to opt for risk alleviation, i.e., mitigation, by the use of incentives. The certainty of attack depends on the fact that the Customers do not receive any advantage from the reduction of the estimate error, to which they contribute.
A strategy for mitigating the risk of attack consists in transferring part of benefit from the cost reduction from the Vendor to the Customers. The Vendor should enact a plan of incentives, e.g. taking the form of discounts on the merchandise.

The Strict fairness-oriented incentive schema. The simplest incentive based schema consists in the Vendor giving each actor his Shapley Value, i.e., keeping half of the benefit for himself and dividing the reminder half proportionally among the Customers cased on each one's estimated demand volume. This zeroes the $v_i^{Shapley} - v_i^{factual}$ and brings to zero the risk related to perceived unfairness.

Partially fair incentive schemas. If the vendor does not want to loose half of the current advantage, he could opt for dividing with the other actors a quota q less than the fair 50% of the revenue. In this case in principle he could either share that quota q proportionally among the Customers, or give to the Customers with largest demand volume their Shapley Value dictated share until q is exhausted. Typically – due to the functional form of the perceived unfairness to attack probability function – the latter solution grants a better trade-off to the Vendor.
Indeed, the Customers with the highest impact are those whose probability of attack should be reduced first. After establishing a budget , i.e., a portion of the cost reduction, the Vendor proposes to all Customers discounts linked to the performance of the overall forecasting. Through private deals with

top ranked Clients, the Vendor should grant higher proportions of the incentive budget to the top ranked Customers.
Which incentive schema works better can be determined by direct computation, by means of the risk assessment tool.

Chapter 6

Tool Description

6.1 Introduction

In this chapter we describe our open source web tool that supports the methodology for risk assessment previously described. The tool supports both the modeling and the simulation of business processes executed on the cloud. The architecture of the tool has been completely redesigned and adopts a client-server model overcoming some of the disadvantages of the first prototypical version.

6.2 Simulator Editor

The editor we developed has the primary aim to support the process of representing and simulating cloud-based business processes for risk analysis purposes, enforcing at the same time some compliance and consistency checks. The procedure to implement the protocol starts with the creation of a graph, where the nodes represent the actors participating in the process, modeled according to the Bogdanov model [7], while the edges are the exchanged messages.

Once the model has been defined, a step-by-step simulation can be started in order to analyse the information flow among the actors. Our editor allows to:

- Create, import and export processes.
- Check the consistency of the process against the constraints imposed by the simulated encryption techniques. For example, in the model only actors IN and $COMP$ are allowed to send data and the management of the correct number of shares in a system.
- Compute the risk of collusion-to-misbehave for each subset of actors considering external input.
- Display charts and graphs to help measure the stepwise risk of disclosure during the execution of the protocol.

Our editor represents processes as graphs. At the implementation level, the editor takes as input a graph $G = (V, E)$ containing a set of edges E, representing messages exchanged, such that each edge is assigned a time instant from 0 to t, and a set of nodes V, representing three categories of actors:

- IN - Actors who collect and send data to the process.
- $COMP$ - Actors that perform a computation in response to the input data of the IN actors.

- *RES* - Actors who receive and display the result of the computation performed by the *COMP* actors.

Once the input graph has been stored, the editor verifies the following constraints:

1. The *IN* actors cannot have input edges (i.e., if $v \in IN$ then $uv \notin E$ for any $v \in V$).
2. The *RES* actors (nodes) belonging to V cannot have output edges (i.e., if $u \in RES$ then $uv \notin E$ for any $v \in V$).

This way, only *IN* and *COMP* actors (nodes) have fan-out output edges.

6.3 Architecture design

With respect to the original design described in deliverable 31.2, the Risk Assessment tool has been redesigned toward a client-server architecture. The original version of the tool was a client-side application, consisting of a simple script running within the client browser. This way, the computational power and hardware resources available depended on the client alone and could not be quantified in advance. The Client-server design structured the tool as a pair of cooperating programs within the application, using the RESTful paradigm for service invocation. The server component provides a risk assessment service to one or many clients, which initiate requests for such services. In other words, the server receives the business process representation and performs the risk analysis, while the client receives the server's analysis output in a character-oriented format (json in the current implementation) and supports its graphical display. More in detail, the core of the REST service is the Post /SVSupp primitive that receives the json data of the process model as input and computes the Shapley Values of the process participants and the probability of malicious behaviour. At each step of the business process the server computes for all participants:

- the information held by each, the one accessible to the cloud provider and their values
- the probability of each participant to misbehave by performing the Shapley value analysis

With respect to the original version of the application described in Deliverable 31.2, the tool now handles all computations taking into account the cloud provider. The return value sent to the client is the Json encoding of the updated model plus a table that, for each supplier, specifies the Shapley value and the probability of malicious behavior. This clean separation of duties between client and server makes it possible to develop alternative clients capable of handling server output in different format (graphs, cvs data etc).

6.3.1 Implementation Notes

The server side of our application is a Web application written in Java, relying on Spark libraries for the web functionalities and on Jackson libraries for parsing the json representation of the business process received from the client. The main classes include:

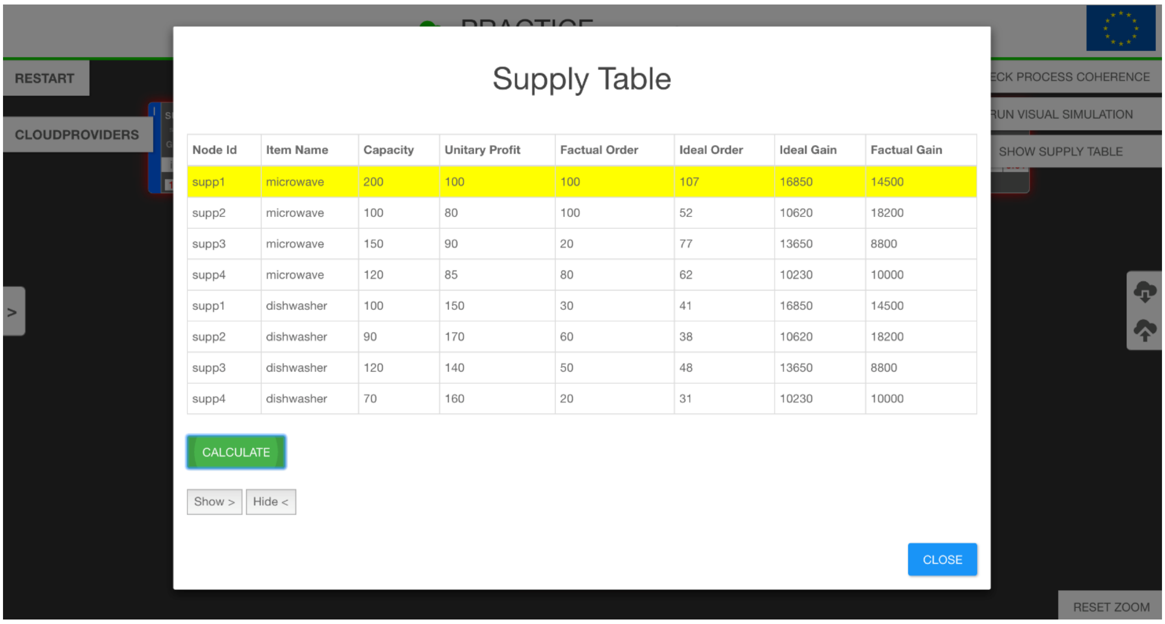
1. A class Actors that is the root of a hierarchy modeling all actors who can be involved in the business process. There are two specializations of that class: Nodes and Providers (the latter modeling the cloud provider and its employees).
2. The ShapleyValues class that encapsulates the algorithm's need for Shapley value analysis, including a permutation generator.

3. A collusion set generator that computes for every possible coalition the power set, its value and its probability.

In turn, the client side of the application is written in JavaScript. For displaying the process under analysis and the results received from the server, the client uses the standard SVG vector graphics library and the Canvas HTML extension. The client-side functionalities have been implemented using the following languages and libraries:

1. *JavaScript*. The Web application is fully implemented through the use of JavaScript. Some precautions and optimizations have been used to obtain a satisfactory product, overcoming the limitations of the selected language.
2. *AngularJS*. It is a framework that allows the creation of web applications using MVC client-side. The choice has been motivated by the possibility to manage dynamic and organization projects in medium / large as the following ones. This framework also helped to break down the part of the inherent logic of the simulation with the layout of the process through the use of the properties of two-way data binding where the data model and controller are linked by an internal loop of AngularJS. By doing so, a modification to the model is instantly passed to the controller and vice-versa.
3. *SVG*. Scalable Vector Graphics is a technology that can display objects of vector graphics through the XML specification. This technology has allowed, along with the library d3.js, to visualize the graph used for simulation, allowing smooth and easy visual interpretation.
4. *Canvas*. Canvas is an extension of HTML standard that allows dynamic rendering of bitmap images manageable through JavaScript. By using the CanvasJS library we can visualize the probability of collusion between all possible subsets of actors involved in the process.
5. *Bootstrap*. Bootstrap allowed to give a standard layout to the project, also allowing aresponsive design which supports multiple resolutions and devices.

Fig .6.1 shows the table returned by the server.



Supply Table

Node Id	Item Name	Capacity	Unitary Profit	Factual Order	Ideal Order	Ideal Gain	Factual Gain
supp1	microwave	200	100	100	107	16850	14500
supp2	microwave	100	80	100	52	10620	18200
supp3	microwave	150	90	20	77	13650	8800
supp4	microwave	120	85	80	62	10230	10000
supp1	dishwasher	100	150	30	41	16850	14500
supp2	dishwasher	90	170	60	38	10620	18200
supp3	dishwasher	120	140	50	48	13650	8800
supp4	dishwasher	70	160	20	31	10230	10000

Figure 6.1: Supply table created by client

6.3.2 A Sample Execution

The following screenshots demonstrate the Shapley value analysis of a supply chain business process using our tool. In Figure 6.2 the main interface is shown. Through the buttons on the right side, the user can upload a JSON representation of a previously created project and download it through the second button. Figure 6.3 shows how it is possible to create a new process through the addition of nodes and arcs which represent, respectively, actors and communications.
After the Shapley Value computation, (Fig. 6.4) the screenshots of Fig. 6.5 and of 6.6 show the probabilities of malicious behavior for each participant at step 1 and 2 of the business process execution.

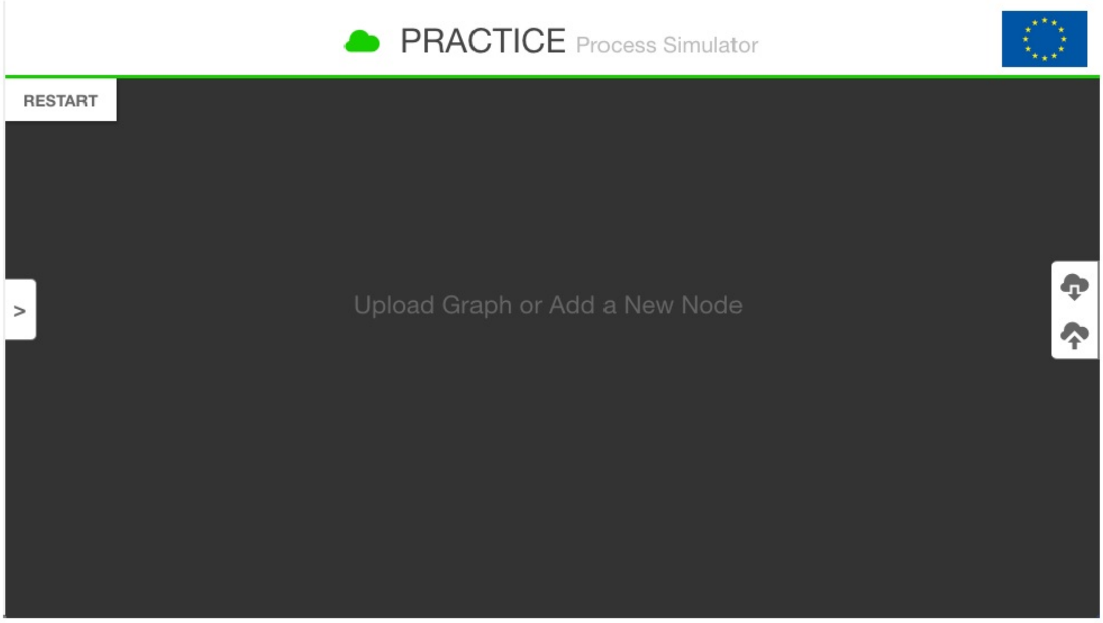
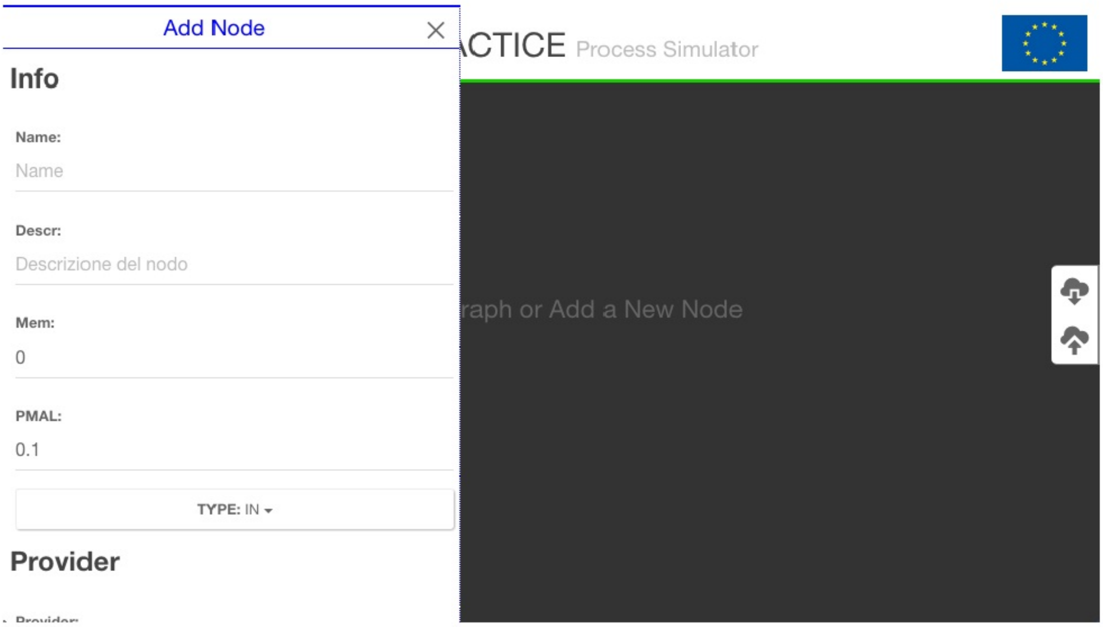


Figure 6.2: Main page, controlled by the MAIN controller



Add Node

Info

Name: Name

Descr: Descrizione del nodo

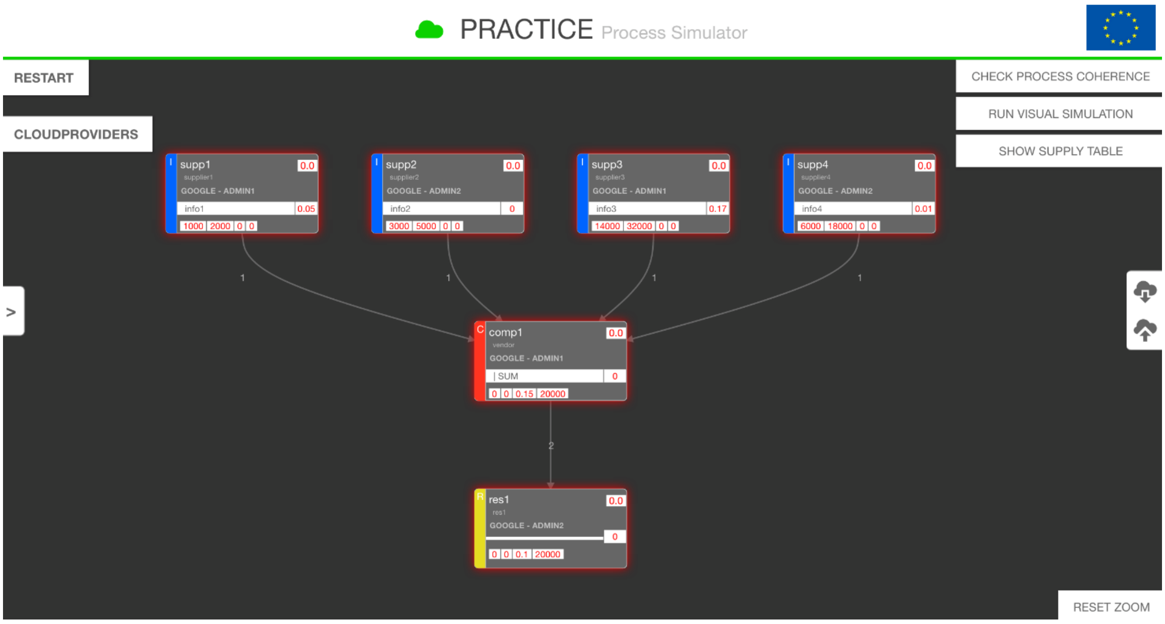
Mem: 0

PMAL: 0.1

TYPE: IN

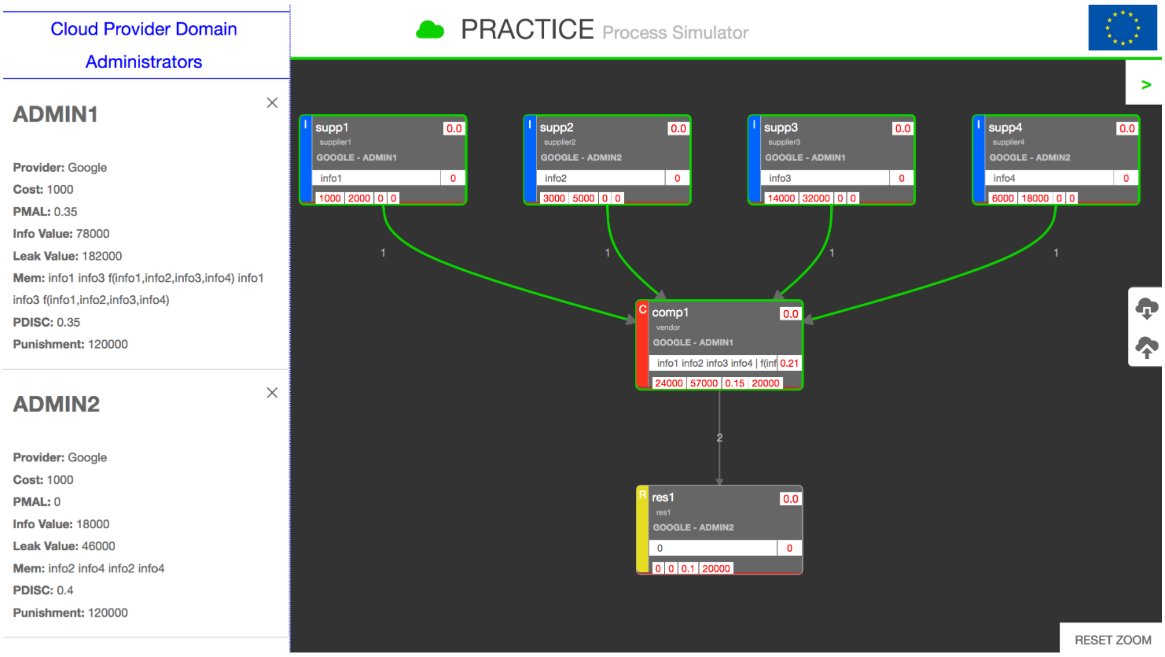
Provider

Figure 6.3: Main Page with overlay used to add nodes



PRACTICE Process Simulator
RESTART
CLOUDPROVIDERS
CHECK PROCESS COHERENCE
RUN VISUAL SIMULATION
SHOW SUPPLY TABLE
I supp1 0.0 supplier1 GOOGLE - ADMIN1 info1 0.05 1000 2000 0 0
I supp2 0.0 supplier2 GOOGLE - ADMIN2 info2 0 3000 5000 0 0
I supp3 0.0 supplier3 GOOGLE - ADMIN1 info3 0.17 14000 32000 0 0
I supp4 0.0 supplier4 GOOGLE - ADMIN2 info4 0.01 6000 18000 0 0
1
1
1
1
C comp1 0.0 vendor GOOGLE - ADMIN1 | SUM 0 0 0 0.15 20000
2
R res1 0.0 res1 GOOGLE - ADMIN2 0 0 0 0.1 20000
>
RESET ZOOM

Figure 6.4: Probability after Shapley value computation.


Cloud Provider Domain Administrators

ADMIN1

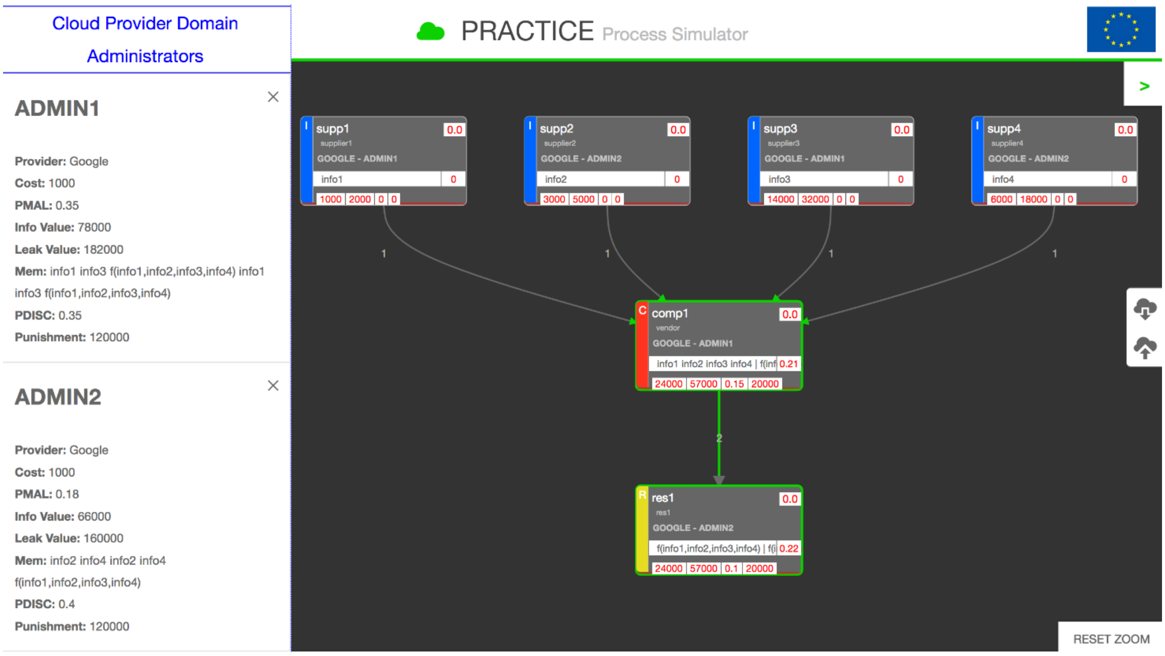
Provider: Google
Cost: 1000
PMAL: 0.35
Info Value: 78000
Leak Value: 182000
Mem: info1 info3 f(info1,info2,info3,info4) info1 info3 f(info1,info2,info3,info4)
PDISC: 0.35
Punishment: 120000

ADMIN2

Provider: Google
Cost: 1000
PMAL: 0
Info Value: 18000
Leak Value: 46000
Mem: info2 info4 info2 info4
PDISC: 0.4
Punishment: 120000

PRACTICE Process Simulator
>
supp1 supplier1 GOOGLE - ADMIN1 0.0
info1 0
1000 2000 0 0
supp2 supplier2 GOOGLE - ADMIN2 0.0
info2 0
3000 5000 0 0
supp3 supplier3 GOOGLE - ADMIN1 0.0
info3 0
14000 32000 0 0
supp4 supplier4 GOOGLE - ADMIN2 0.0
info4 0
6000 18000 0 0
1
1
1
1
comp1 vendor GOOGLE - ADMIN1 0.0
info1 info2 info3 info4 | f(inf 0.21
24000 57000 0.15 20000
2
res1 res1 GOOGLE - ADMIN2 0.0
0 0
0 0 0.1 20000
RESET ZOOM

Figure 6.5: Pmal at step 1 of the business process


Cloud Provider Domain Administrators

ADMIN1

Provider: Google
Cost: 1000
PMAL: 0.35
Info Value: 78000
Leak Value: 182000
Mem: info1 info3 f(info1,info2,info3,info4) info1 info3 f(info1,info2,info3,info4)
PDISC: 0.35
Punishment: 120000

ADMIN2

Provider: Google
Cost: 1000
PMAL: 0.18
Info Value: 66000
Leak Value: 160000
Mem: info2 info4 info2 info4 f(info1,info2,info3,info4)
PDISC: 0.4
Punishment: 120000

PRACTICE Process Simulator
>
supp1 supplier1 GOOGLE - ADMIN1 0.0
info1 0
1000 2000 0 0
supp2 supplier2 GOOGLE - ADMIN2 0.0
info2 0
3000 5000 0 0
supp3 supplier3 GOOGLE - ADMIN1 0.0
info3 0
14000 32000 0 0
supp4 supplier4 GOOGLE - ADMIN2 0.0
info4 0
6000 18000 0 0
1
1
1
1
comp1 vendor GOOGLE - ADMIN1 0.0
info1 info2 info3 info4 | f(inf 0.21
24000 57000 0.15 20000
2
res1 res1 GOOGLE - ADMIN2 0.0
f(info1,info2,info3,info4) | f(i 0.22
24000 57000 0.1 20000
RESET ZOOM

Figure 6.6: Pmal at step 2 of the business process.

6.4 Case Study

In this section we show a numerical example of Shapley Value computation and the correlated probability of misrepresentation attack by the customers considering the use case examined in section 5.4. where the Arcelik supply chain is modeled and analyzed.

The model is defined in the tool with nodes (Customers and Vendor) and edges (from Customer to Vendor) and the data relative to the supply chain are inserted in a table where every row contains:

NodeId of the customer, the SKU code of the item, the supply quantity (derived from historical supply data volumes), the unitary price, the total price (quantity * unitary price), the discount in percentage on the original price and the discounted price, the Shapley value and the probability of attack. The last two fields are computed by the server.
In the use case it is assumed that the collaborative forecast method COLFOR-* produces a cost reduction for the Supply Chain of 10%. According with the model described in sec. 5.4, the Shapley Value of Vendor is 50%, while the remainder 50% is partitioned among the customers, proportionally to their supply quantity. This means, as a consequence of the fact that the cost reduction is 10%, that the Vendor should withhold the 5% of the cost reduction, while the remaining 5% is shared among the customers e.g. in the form of a discount. Since the Customers share the discount proportionally to their importance (their order volumes), the ideal discount should, trivially be the 5% of their order.
The following screen shots show the graph of the model and the tables of the supply chain with three different discount scenarios before and after server computation:

1. No discount scenario (Figures 6.7,–6.9)

2. Shapley Value based discount scenario (Figures 6.10–6.11)

3. Larger discount for larger volumes scenario (Figures 6.12–6.13)

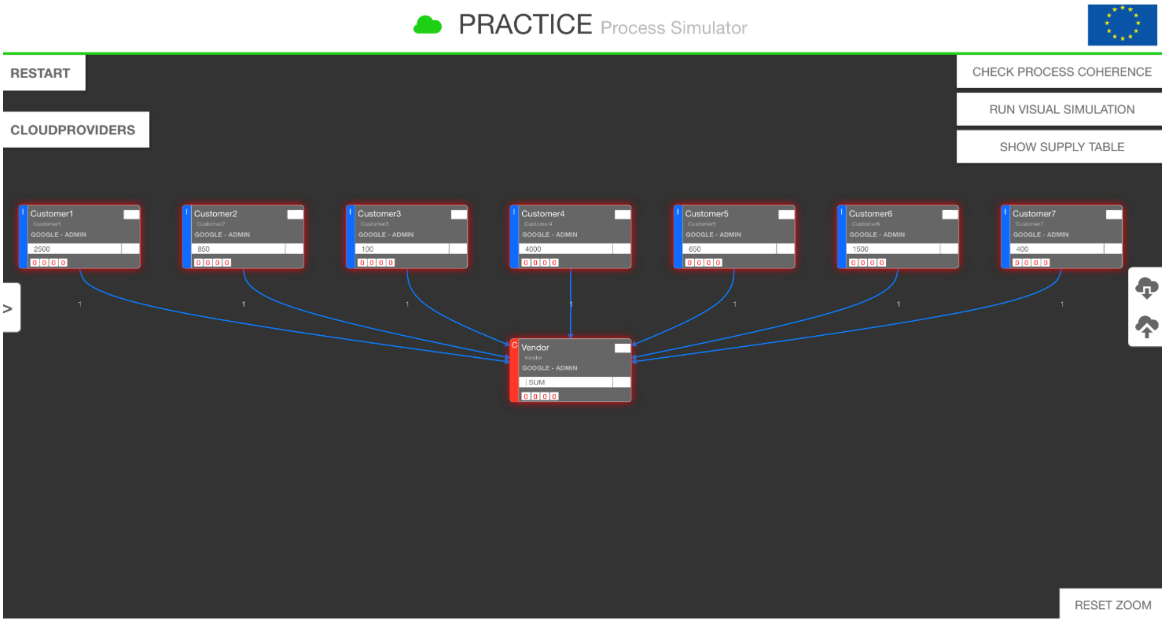
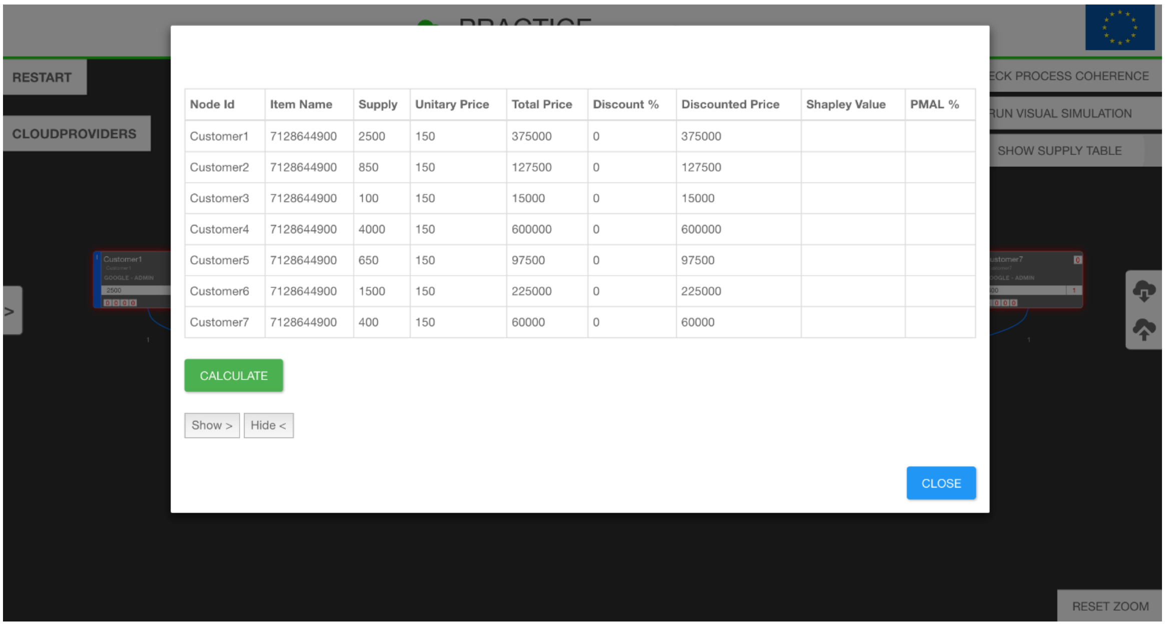
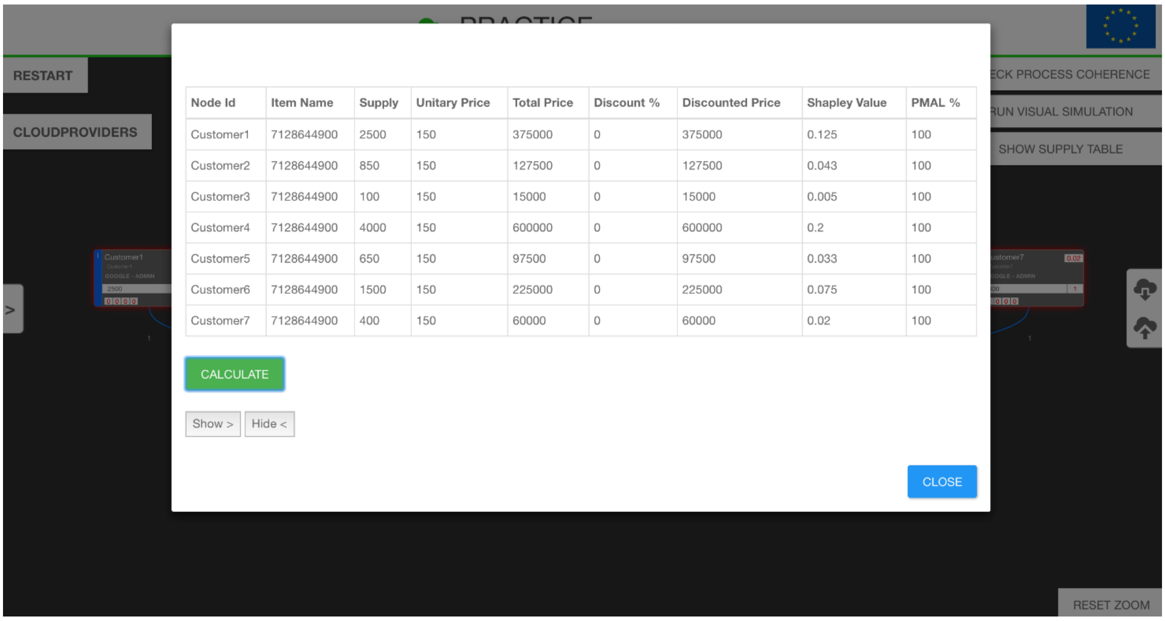


Figure 6.7: Initial model



Node Id	Item Name	Supply	Unitary Price	Total Price	Discount %	Discounted Price	Shapley Value	PMAL %
Customer1	7128644900	2500	150	375000	0	375000		
Customer2	7128644900	850	150	127500	0	127500		
Customer3	7128644900	100	150	15000	0	15000		
Customer4	7128644900	4000	150	600000	0	600000		
Customer5	7128644900	650	150	97500	0	97500		
Customer6	7128644900	1500	150	225000	0	225000		
Customer7	7128644900	400	150	60000	0	60000		

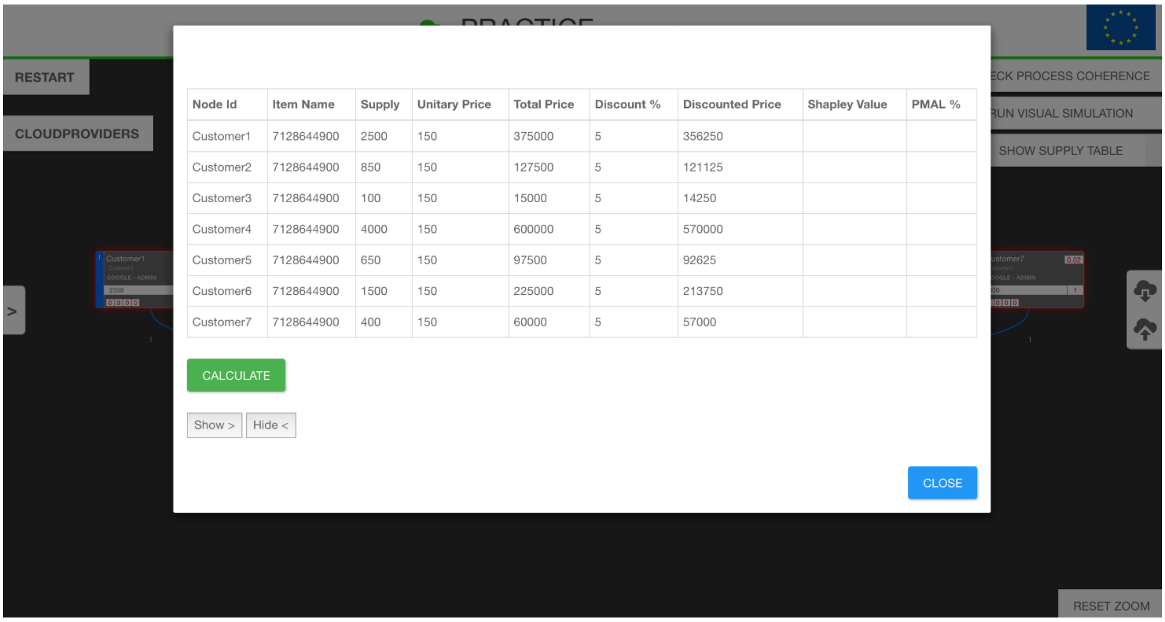
Figure 6.8: First input with no discount. Shapley Value, perceived unfairness index (not shown) and probability of malicious behavior (PMAL %) are still to be computed.



Node Id	Item Name	Supply	Unitary Price	Total Price	Discount %	Discounted Price	Shapley Value	PMAL %
Customer1	7128644900	2500	150	375000	0	375000	0.125	100
Customer2	7128644900	850	150	127500	0	127500	0.043	100
Customer3	7128644900	100	150	15000	0	15000	0.005	100
Customer4	7128644900	4000	150	600000	0	600000	0.2	100
Customer5	7128644900	650	150	97500	0	97500	0.033	100
Customer6	7128644900	1500	150	225000	0	225000	0.075	100
Customer7	7128644900	400	150	60000	0	60000	0.02	100

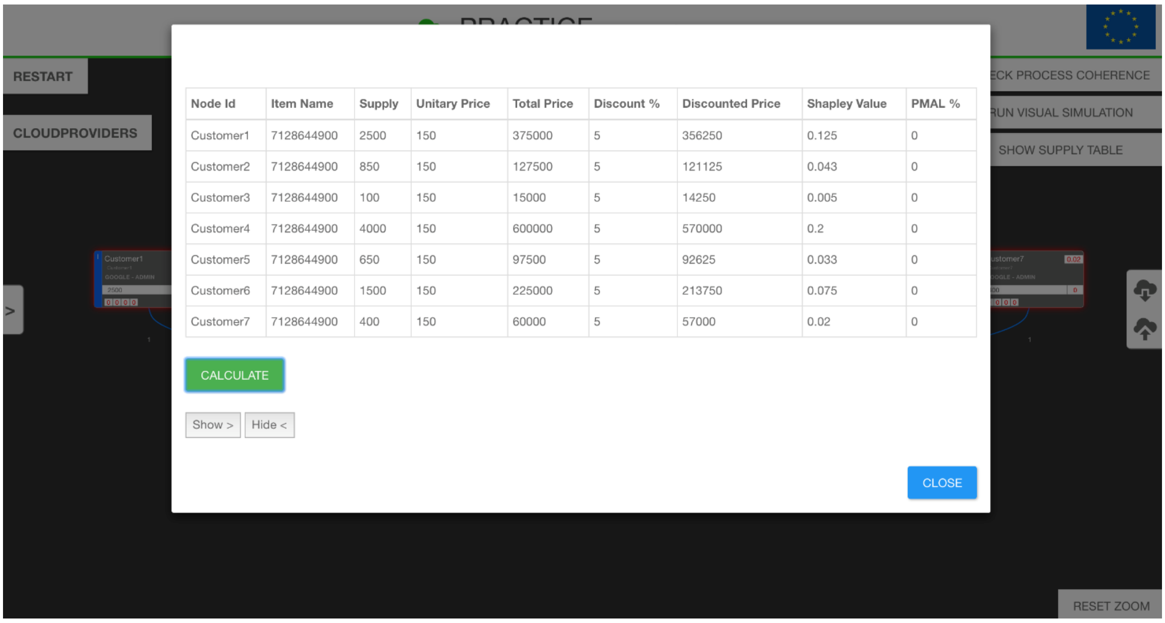
Figure 6.9: Table filled with data from server computation. Notice then PMAL is 1 for all customers (see text).

No discount scenario. Figure 6.7 shows the initial status of the computation, Figure 6.8 the data provided in input: notice that the discount provided to all the Customers is zero. 6.9 shows the computed Shapley Value, that is always positive. Being the payoff to the Customer (under the form of a discount) equal to zero, numerator and denominator of the perceived unfairness index ϕ are always equal; since ϕ is always 1, also the probability p_i of malicious behavior by an actor i, indicated in the figure by PMAL, is always 1.



Node Id	Item Name	Supply	Unitary Price	Total Price	Discount %	Discounted Price	Shapley Value	PMAL %
Customer1	7128644900	2500	150	375000	5	356250		
Customer2	7128644900	850	150	127500	5	121125		
Customer3	7128644900	100	150	15000	5	14250		
Customer4	7128644900	4000	150	600000	5	570000		
Customer5	7128644900	650	150	97500	5	92625		
Customer6	7128644900	1500	150	225000	5	213750		
Customer7	7128644900	400	150	60000	5	57000		

Figure 6.10: Second input with 5% discount

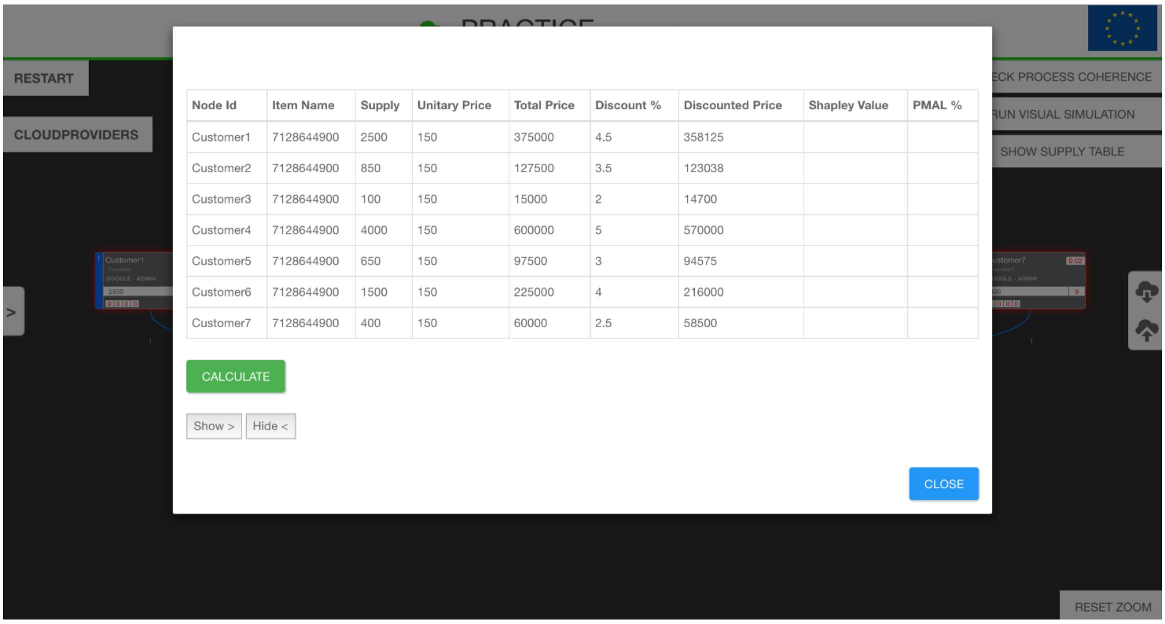


Node Id	Item Name	Supply	Unitary Price	Total Price	Discount %	Discounted Price	Shapley Value	PMAL %
Customer1	7128644900	2500	150	375000	5	356250	0.125	0
Customer2	7128644900	850	150	127500	5	121125	0.043	0
Customer3	7128644900	100	150	15000	5	14250	0.005	0
Customer4	7128644900	4000	150	600000	5	570000	0.2	0
Customer5	7128644900	650	150	97500	5	92625	0.033	0
Customer6	7128644900	1500	150	225000	5	213750	0.075	0
Customer7	7128644900	400	150	60000	5	57000	0.02	0

Figure 6.11: Table filled with data from server computation. Notice then PMAL is 0 for all customers because the discount is the ideal one.

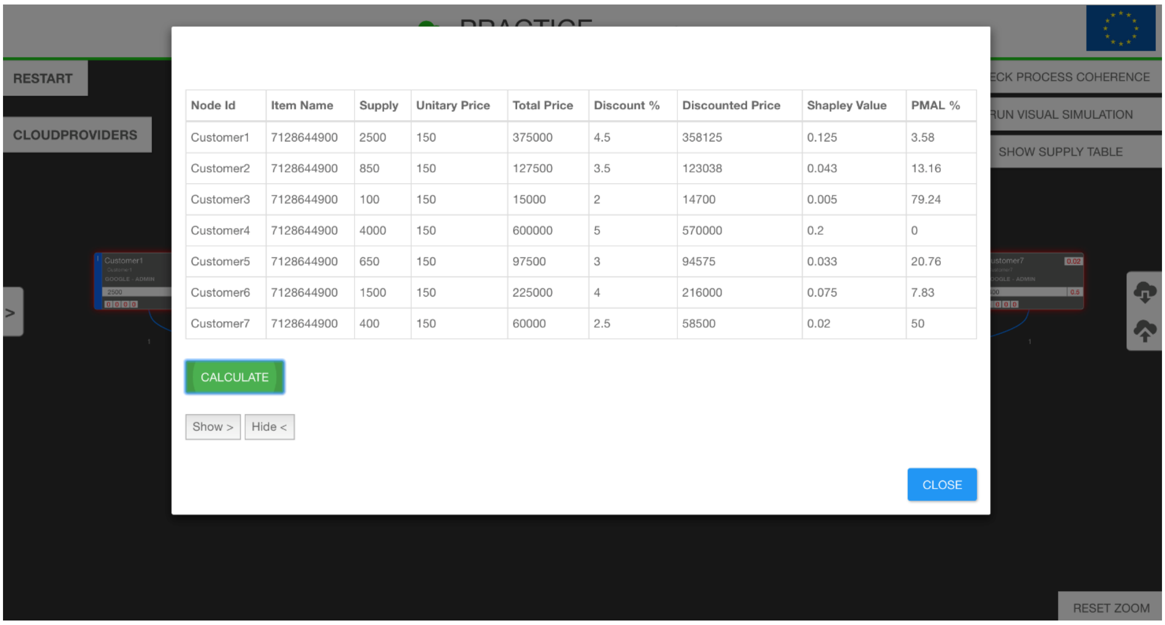
Shapley Value based discount scenario. Figure 6.10 shows the data provided in input: notice that the discount provided to all the Customers the 5% of their supply volume. As discussed above this discount reflects a Shapley Value based division. Figure 6.11 shows the computed Shapley Value: notice that the Customers have different the Shapley Values, because their volumes are different, and that also the absolute values of the discounts are different, however, the discounts are proportional to the Shapley Value. Being the payoff to the Customer (under the form of a discount) proportional to the Shapley Value, the numerator of the perceived unfairness index ϕ is zero; since ϕ is always 0, also the probability p_i of malicious behavior by an actor i, indicated in the figure by PMAL, is always 0

Larger discount for larger volumes. Figure 6.12 shows the data provided in input in a scenario where the larger volume Customers are given larger discounts. Figure 6.13 shows the computed Shapley Value and PMAL: Customers with discount percentage close to their Shapley Value are less kin to attack.



Node Id	Item Name	Supply	Unitary Price	Total Price	Discount %	Discounted Price	Shapley Value	PMAL %
Customer1	7128644900	2500	150	375000	4.5	358125		
Customer2	7128644900	850	150	127500	3.5	123038		
Customer3	7128644900	100	150	15000	2	14700		
Customer4	7128644900	4000	150	600000	5	570000		
Customer5	7128644900	650	150	97500	3	94575		
Customer6	7128644900	1500	150	225000	4	216000		
Customer7	7128644900	400	150	60000	2.5	58500		

Figure 6.12: Third input with variable discount increasing with the importance of customers



Node Id	Item Name	Supply	Unitary Price	Total Price	Discount %	Discounted Price	Shapley Value	PMAL %
Customer1	7128644900	2500	150	375000	4.5	358125	0.125	3.58
Customer2	7128644900	850	150	127500	3.5	123038	0.043	13.16
Customer3	7128644900	100	150	15000	2	14700	0.005	79.24
Customer4	7128644900	4000	150	600000	5	570000	0.2	0
Customer5	7128644900	650	150	97500	3	94575	0.033	20.76
Customer6	7128644900	1500	150	225000	4	216000	0.075	7.83
Customer7	7128644900	400	150	60000	2.5	58500	0.02	50

Figure 6.13: Table filled with data from server computation. Notice then PMAL increases with the decrease of the percentage discount.

Bibliography

[1] https://www.enisa.europa.eu/topics/threat-risk-management/risk-management/current-risk/risk-management-inventory/rm-process/risk-treatment.

[2] http://www.lamsade.dauphine.fr/ airiau/teaching/index.html.

[3] Robert J Aumann and Jacques H Dreze. Cooperative games with coalition structures. *International Journal of game theory*, 3(4):217–237, 1974.

[4] Samik Basu and Tevfik Bultan. Choreography conformance via synchronizability. In Sadagopan Srinivasan, Krithi Ramamritham, Arun Kumar, M. P. Ravindra, Elisa Bertino, and Ravi Kumar, editors, *Proceedings of the 20th International Conference on World Wide Web, WWW 2011, Hyderabad, India, March 28 - April 1, 2011*, pages 795–804. ACM, 2011.

[5] Muli Ben-Yehuda, Michael D. Day, Zvi Dubitzky, Michael Factor, Nadav Har'El, Abel Gordon, Anthony Liguori, Orit Wasserman, and Ben-Ami Yassour. The turtles project: Design and implementation of nested virtualization. In *Proceedings of the 9th USENIX Conference on Operating Systems Design and Implementation*, OSDI'10, pages 1–6, Berkeley, CA, USA, 2010. USENIX Association.

[6] Matt Bishop, Heather M Conboy, Huong Phan, Borislava I Simidchieva, George S Avrunin, Lori A Clarke, Leon J Osterweil, and Sean Peisert. Insider threat identification by process analysis. In *Security and Privacy Workshops (SPW), 2014 IEEE*, pages 251–264. IEEE, 2014.

[7] Dan Bogdanov, Liina Kamm, Sven Laur, and Pille Pruulmann-Vengerfeldt. Secure multi-party data analysis: end user validation and practical experiments. Cryptology ePrint Archive, Report 2013/826, 2013.

[8] Michael Burrows, Martin Abadi, and Roger Needham. A logic of authentication. *ACM Trans. Comput. Syst.*, 8(1):18–36, February 1990.

[9] Shakeel Butt, H. Andrés Lagar-Cavilla, Abhinav Srivastava, and Vinod Ganapathy. Self-service cloud computing. In *Proceedings of the 2012 ACM Conference on Computer and Communications Security*, CCS '12, pages 253–264, New York, NY, USA, 2012. ACM.

[10] Cas J. Cremers. The scyther tool: Verification, falsification, and analysis of security protocols. In *Proceedings of the 20th International Conference on Computer Aided Verification*, CAV '08, pages 414–418, Berlin, Heidelberg, 2008. Springer-Verlag.

[11] Theo Driessen. Cooperative games, solutions and applications. series c: Game theory. *Mathematical Programming and Operations Research. Kluwer Academic Publishers, Dordrecht, The Netherlands*, 1988.

[12] F. Javier Thayer Fábrega. Strand spaces: Proving security protocols correct. *J. Comput. Secur.*, 7(2-3):191–230, March 1999.

[13] Ronald Fagin, Yoram Moses, Joseph Y Halpern, and Moshe Y Vardi. *Reasoning about knowledge*. MIT press, 2003.

[14] Ulrich Faigle, Walter Kern, and Jeroen Kuipers. On the computation of the nucleolus of a cooperative game. *International Journal of Game Theory*, 30(1):79–98, 2001.

[15] DB Gillies. Some theorems on n-person games. princeton university. *Unpublished doctoral dissertation.)[aAMC]*, 1953.

[16] Aaron Hunter and James P. Delgrande. Belief change and cryptographic protocol verification. In *Proceedings of the 22Nd National Conference on Artificial Intelligence - Volume 1*, AAAI'07, pages 427–433. AAAI Press, 2007.

[17] Andy Jones and Debi Ashenden. *Risk management for computer security: Protecting your network & information assets*. Butterworth-Heinemann, 2005.

[18] Ehud Kalai and Dov Samet. On weighted shapley values. *International Journal of Game Theory*, 16(3):205–222, 1987.

[19] Eric Keller, Jakub Szefer, Jennifer Rexford, and Ruby B. Lee. Nohype: Virtualized cloud infrastructure without the virtualization. *SIGARCH Comput. Archit. News*, 38(3):350–361, June 2010.

[20] Apurva Kumar. A belief logic for analyzing security of web protocols. In *Proceedings of the 5th International Conference on Trust and Trustworthy Computing*, TRUST'12, pages 239–254, Berlin, Heidelberg, 2012. Springer-Verlag.

[21] Fabio Massacci. Automated reasoning and the verification of security protocols. In *Proceedings of the International Conference on Automated Reasoning with Analytic Tableaux and Related Methods*, TABLEAUX '99, pages 32–33, London, UK, UK, 1999. Springer-Verlag.

[22] Roger B Myerson. Conference structures and fair allocation rules. *International Journal of Game Theory*, 9(3):169–182, 1980.

[23] Jason RC Nurse, Oliver Buckley, Philip A Legg, Michael Goldsmith, Sadie Creese, Gordon RT Wright, and Monica Whitty. Understanding insider threat: A framework for characterising attacks. In *Security and Privacy Workshops (SPW), 2014 IEEE*, pages 214–228. IEEE, 2014.

[24] Guillermo Owen. Game theory academic press. *San Diego*, 1995.

[25] Bezalel Peleg and Peter Sudhölter. *Introduction to the theory of cooperative games*, volume 34. Springer Science & Business Media, 2007.

[26] August-Wilhelm Scheer and Markus Nüttgens. ARIS architecture and reference models for business process management. In Wil M. P. van der Aalst, Jörg Desel, and Andreas Oberweis, editors, *Business Process Management, Models, Techniques, and Empirical Studies*, volume 1806 of *Lecture Notes in Computer Science*, pages 376–389. Springer, 2000.

[27] David Schmeidler. The nucleolus of a characteristic function game. *SIAM Journal on applied mathematics*, 17(6):1163–1170, 1969.

[28] Lloyd S Shapley. *Additive and non-additive set functions*. Princeton University, 1953.

[29] Udo Steinberg and Bernhard Kauer. Nova: A microhypervisor-based secure virtualization architecture. In *Proceedings of the 5th European Conference on Computer Systems*, EuroSys '10, pages 209–222, New York, NY, USA, 2010. ACM.

[30] Paul F. Syverson. Knowledge, belief, and semantics in the analysis of cryptographic protocols. *J. Comput. Secur.*, 1(3-4):317–334, May 1992.

[31] Paul F. Syverson. Adding time to a logic of authentication. In *Proceedings of the 1st ACM Conference on Computer and Communications Security*, CCS '93, pages 97–101, New York, NY, USA, 1993. ACM.

[32] Fengzhe Zhang, Jin Chen, Haibo Chen, and Binyu Zang. Cloudvisor: Retrofitting protection of virtual machines in multi-tenant cloud with nested virtualization. In *Proceedings of the Twenty-Third ACM Symposium on Operating Systems Principles*, SOSP '11, pages 203–216, New York, NY, USA, 2011. ACM.