D31.1

Risk assessment and current legal status on data protection

Project number:	609611
Project acronym:	**PRACTICE**
Project title:	Privacy-Preserving Computation in the Cloud
Project Start Date:	1 November, 2013
Duration:	36 months
Programme:	FP7/2007-2013
Deliverable Type:	Report
Reference Number:	ICT-609611 / D31.1 / 1.0
Activity and WP:	Activity 3 / WP31
Due Date:	October 2014 - M12
Actual Submission Date:	3rd November, 2014
Responsible Organisation:	UMIL
Editor:	Stelvio Cimato
Dissemination Level:	PU
Revision:	1.0
Abstract:	This deliverable reports on the current legal framework regulating the storage and processing the data on the cloud and introduces a risk assessment methodology to analyze the business risks associated with outsourcing data.
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

Stelvio Cimato (UMIL)

Contributors (ordered according to beneficiary numbers)

Ernesto Damiani (UMIL)
Valerio Bellandi (UMIL)
Stelvio Cimato (UMIL)
Gabriele Gianini (UMIL)
Gerald Spindler (UGOE)
Matthis Grenzer (UGOE)
Christopher Schwanitz (UGOE)
David Koppe (UGOE)
Niklas Heitmüller (UGOE)
Sonja Hagenhoff (UGOE)
Tim Kostka (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. 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 outline the state of the art in assessing and managing risks for cloud-based business processes. Then, we describe the first version of a process-oriented 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 multi-party business process. It enables comparing risks of disclosure "before" and "after" the deployment of PRACTICE security controls. Also, our methodology is designed to be pluggable into existing qualitative and semi-quantitative risk management frameworks. As a proof-of-concept, in this deliverable our quantitative risk assessment methodology is applied to the analysis of business processes implementing Vickrey auctions, i.e. auctions where the bidder putting forward the second-best offer is successful.

Contents

Chapter 1

Introduction

Work package 31's objectives include the evaluation of the legal aspects related to the outsourcing of data and of computation to the cloud, and the development of models and techniques to quantify the business risks associated with data sharing in collaborative services. The work package's long term goals are (i) the clarification of the legal framework regulating the placing and the processing of sensitive data in locations where different privacy regulations hold, possibly establishing a set of guidelines compliant to international legal frameworks (ii) the development of a quantitative risk assessment methodology for the deployment of secure computation protocols on the cloud. The latter should support eventually the computation of Return On Investment (ROI) in security controls when data and process computation are (jointly or separately) outsourced. Also, it should allow the computation and visualization of *risk profiles* showing risk concerning different parties in different scenarios.

This deliverable presents the result of the first year of research on these topics. It is divided into two parts. The first part is devoted to a complete overview of the current legal framework regulating data protection in the European Union. In particular, we discuss the EU Data Protection directive currently in force and some other auxiliary regulations, highlighting their relevance to the processing of personal data on the cloud. 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 requirements for the upcoming 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 current legal framework.

The second part of the deliverable is composed of two Chapters. Chapter 4 outlines the state of the art in assessing and managing risks involved in the execution of cloud-based business processes, with special reference to *Privacy Risk Management Framework*s (PRMF). Chapter 5 describes the first version of PRACTICE process-oriented assessment methodology, which can be used to analyze risks that arise when multi-party processes are executed on clouds. Our methodology is based on a simple model for the representation of business processes. Our process model lends itself to visual representation of processes and - hopefully - is close enough to standard business process representation like the *Choreography Description Language* (WS-CDL) to be readily understandable by non-specialists.

In a nutshell, our approach works as follows: first, all actors taking part to the business process under analysis are tagged according to their role; then, for each subset of the actor set (and for each process execution step) we identify the information items that would be reconstructible should the subset's members agree to share all the information they hold at that step. The probability of such sharing to take place is then estimated based on the micro-economics un-

derlying the business process under analysis. Its impact on any other actor taking part to the business process is also estimated via the perceptions of the disclosed information's value on the part of the actor itself.
Our risk assessment methodology supports *comparative analysis* and *quantitative evaluation* of risks related to privacy and confidentiality breaches during the execution of any multi-party business process. Specifically, it enables comparing risks of disclosure of a business process "before" and "after" the deployment within it of PRACTICE security controls. While our methodology's integration interface with standard risk assessment will be fully specified in later deliverables, this document already shows how the methodology was designed to be seamlessly plugged into existing qualitative and semi-quantitative risk management frameworks.
As a proof-of-concept, this deliverable discusses how the methodology can be applied to the analysis of cloud-based business processes implementing auctions. We focus on Vickrey auctions, i.e. auctions where the bidder putting forward the second-best offer is successful. Our description shows how to compute risk involved in a simple version of the Vickrey auction, and how to compare it to the (lower) risk of a version where SMC is used to hide which actor submitted which offer.

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 involved party. Compliance especially with the data protection law can be hard to achieve. 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

Before specific problems for Cloud Computing arising from the data protection law are analysed, the basic functioning and main principles of the current and the upcoming state of law shall be outlined. Only if there is a general understanding of the underlying legal framework can the causes of and possible solutions for the legal difficulties be properly explained and understood.

2.1.1 The Data Protection Directive 95/46/EC

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"). Art. 288 of the Treaty on the Functioning of the European Union defines how the Unions 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. Only if directives state rights

[1] *Hon/Millard/Walden*, Who is Responsible for 'Personal Data' in Cloud Computing?, The Cloud of Unknowing, Part 2, p. 3.

[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`.

for citizens and if they are not implemented in due time by national authorities, citizens may claim those rights directly.

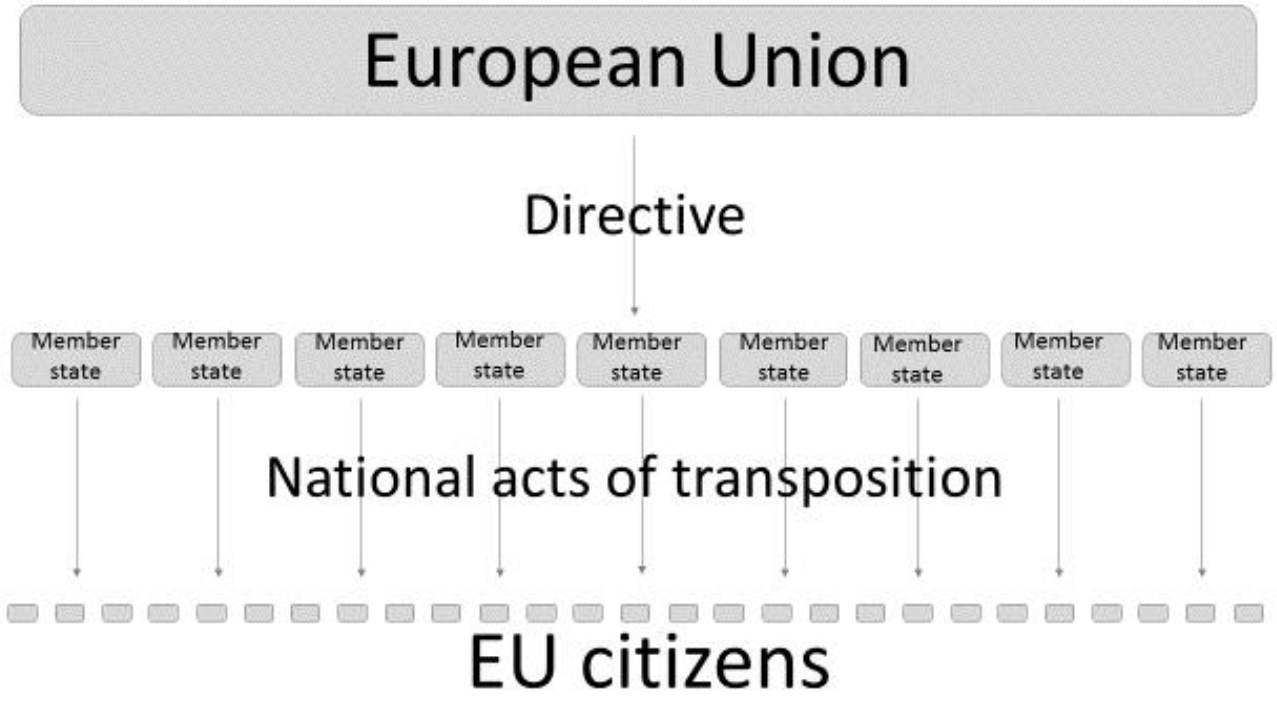

Figure 2.1: EU Directives

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, be it a de minimis harmonization and leaving member states some leeway to pass laws which are going beyond that level, be it a full harmonization preventing member states to go beyond the directive.

Concerning the Directive 95/46/EC the European Court of Justice 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 cannot add new principles relating to the lawfulness of processing or impose additional requirements. [6]

[3] `http://ec.europa.eu/eu_law/introduction/what_directive_en.htm.`

[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 of Luxembourg, available at:`http://ec.europa.eu/anti_fraud/documents/data-protection/dpo/ecj_decisions_relating_data_protection_en.pdf.`

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.

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 2002 Directive 2002/58/EC (Directive on privacy and electronic communications).

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 his activities of the processor as well. Even an end-users machine might be 'equipment situated on the territory of a member state' if it is used for storing a cookie or collecting data with java scripts. [7]

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 crucial where a service is aimed at but, where the resources used for providing this service are located (this principle will change with the upcoming Data Protection Regulation, see 4.1.1). [8] A cloud server in Europe would be qualified as 'equipment' in the sense of the DPD. [9]

If a controller is established on the territory of several Member States, he has to ensure compliance with each of the national laws applicable, Art 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. On the other hand, it is not necessary that the establishment is independent from the controller in order to be considered as a controller itself (for the definition of 'data controller' see 3.6). [10]

[7] As for example stated by the German court KG Berlin in its ruling from 24/01/2014, 5 U 42/12, 28 f., available at:http://www.berlin.de/imperia/md/content/senatsverwaltungen/justiz/kammergericht/presse/5_u_42_12_urteil_vom_24.1.2014_kammergericht_anonymisiert.pdf?start&start&ts=1392399485&file=5_u_42_12_urteil_vom_24.1.2014_kammergericht_anonymisiert.pdf.

[8] *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.

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

[10] 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

One of the cases decided by the ECJ highlights the difficulties in practice to handle the notion of establishment in the DPD. [11] The arguments brought forward by the General Advocate in the case of Google vs. Spain are worth being cited literally in order to highlight the range 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 and, as such, the economic raison d'tre 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." [12]

In the final judgment, the ECJ followed the General Advocates opinion:

of Facebook, also in which European country Facebooks 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'.

[11] ECJ, Judgment from the 13th May 2014 in Case C-131/12 - Google Spain SL/Google Inc. v AEPD/Mario Costeja Gonzalez.

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

> "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." [13]

Applied to cloud computing the DPD will be applicable if the cloud is an entity established within the jurisdiction of a member state, if data processing is carried out in the context of the activities of the establishment. Every instance of processing a provider is about to carry out would then be covered, including the transfer of data to a non EU country. In general, the directive will be applied 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 the provider does not have an establishment in another EU-member state. However, according to the mentioned ECJ decision "Google Spain" it is already sufficient for the application of the directive that there is an establishment of the cloud provider in the EU that fosters the activities of the cloud provider. 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 that from an economic perspective the establishment supports the activities of the cloud provider, for instance like in the Google Spain Case selling of advertisement etc. Hence, it should already be sufficient that an establishment operates the cashing for the cloud provider etc., in order to apply the Data Protection Directive.

2.1.1.2 Material scope and Fundamentals of the DPD

The focus of the directive regarding "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. [14]

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.

[13] ECJ, Judgment from the 13th May 2014 in Casec-131/12 - Google Spain SL/Google Inc. v AEPD/Mario Costeja Gonzalez, rec. 55.

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

2.1.1.2.1 Material scope of the Data Protection Directive

The directive protects only personal data of individuals, 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 the private or job-related area, characteristics, skills of an employee, psychological characteristics or elements of someone's biography. [15]

Whereas the directive applies in general for all kinds of processing data there are still some differences made by the directive. In case of non-automatic processing the directive addresses only data processing stored in a (physical) dossier. However, in this report we will deal only with requirements for automatic processing of data due to the character of cloud computing. Moreover, Article 3 (2) refers to some exceptions: [16]

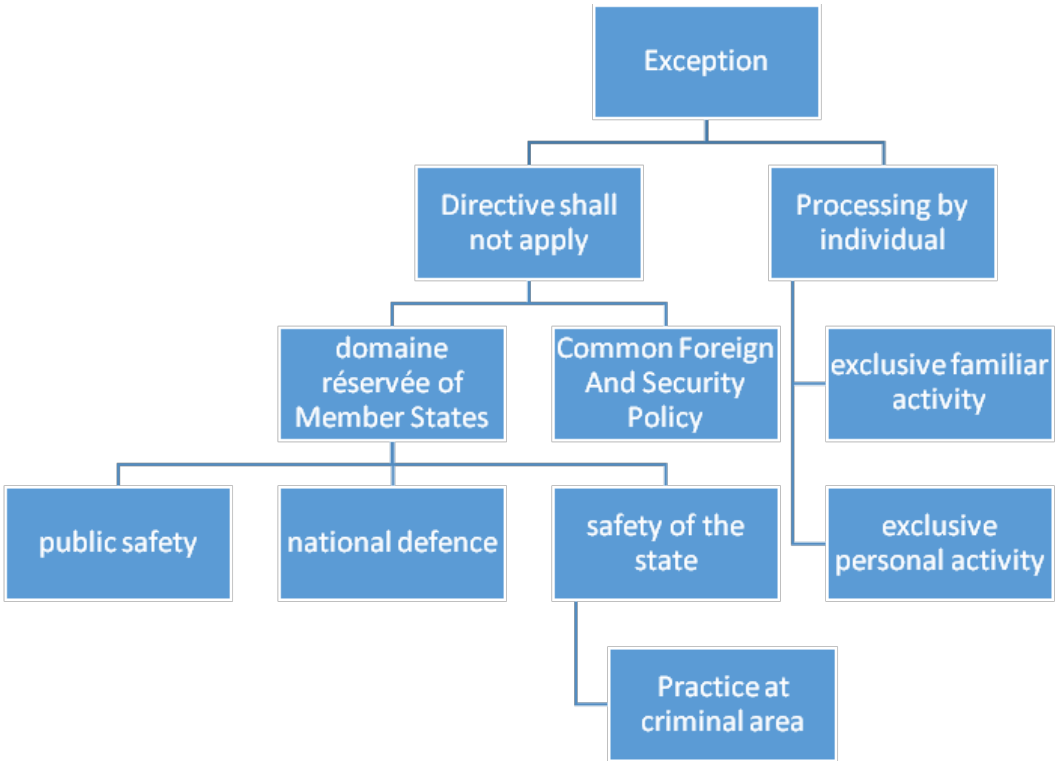

Figure 2.2: Applicability of the Data Protection Directive

One of the exceptions relevant for internet services (and users) refers to exclusive personal and familiar activities which are exempted. Hence, for instance 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 purpose is not a case of the scope of this data protection directive.

[15] *Dammann*, in: Dammann/Simitis, EG-Datenschutzrichtlinie, Art. 2, p.109.
[16] *Ehmann/Helfrich*, EG-Datenschutzrichtlinie, Art. 3, recital 16.

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 Art. 10, 11 The individual has the right to be informed should his personal data be processed. 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. [17] Personal Data can be processed only if the controller complies with the requirements stated in Article 7 and 12. Thus an explicit consent of the data subject is indispensable for the performance of contractual obligations or the entering into a contract.

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 controller to "keep [the data] in a form which permits identification of data subjects for no longer than is necessary for the purposes for 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", Art. 6 (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 Art. 8 (2) are fulfilled such as a specific consent or protecting the vital interests of the data subject etc.

2.1.2 The Proposal for a General Data Protection Regulation

The General Data Protection Regulation is a proposed regulation of the European Union, which harmonizes the rules for the processing of personal data mainly by private companies EU-wide - whereas it is still heavily debated if the Regulation should also touch other sectors such as tax authorities, social security, etc. The outdated [18] Data Protection Directive 95/46/EC from the

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

[18] *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).

year 1995 was intended to encourage the free movement of personal data within Europe by harmonizing national provisions on data protection. [19] The implementation scope of the directive led to different interpretations of the national data protection laws and to a minimum standard. [20] This is to ensure a uniform standard of data protection [21], the protection of personal data and the free movement of such data within the European Union. On October 21st2013 the European Parliaments LIBE Committee (Committee for Civil Liberties, Justice and Home Affairs) adopted a number of proposed changes to the General Data Protection Regulation published by the EU Commission on January 25th 2012. [22] The Home Affairs Committee of the European Parliament launched on October 22nd 2013 the start of negotiations with the European Commission, and the Council of the European Union - the so-called trilogue which is still going on. On March 12th 2014 the European Parliament adopted a legislative resolution on the proposal after the first reading in the parliament, adopting the LIBE Committees changes to the original proposal - being the latest official version. [23] The proposal for a General Data Protection Regulation *maintains the main principle* of the directive 95/46/EC to *generally prohibit the processing* of personal data, unless the person affected has given their consent or legal permission.

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 [24] 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. [...]"

2.1.2.2 Territorial Scope of the GDPR

The territorial scope of the Regulation is specified in three cases, Article 3 Par. 1 - 3: [25]

[19] *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).

[20] *Klar*, ZD 2013, 109 (109 ff.); While one could have understood the Lindqvist- decision of the ECJs (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.

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

[22] 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-Comittees vote, available at: `http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+REPORT+A7-2013-0402+0+DOC+XML+V0//EN`; *Heinemeyer*, Verfahrensstand-Anzeiger; Härting, CR 2013, 715 (715 ff.).

[23] 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`.

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

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

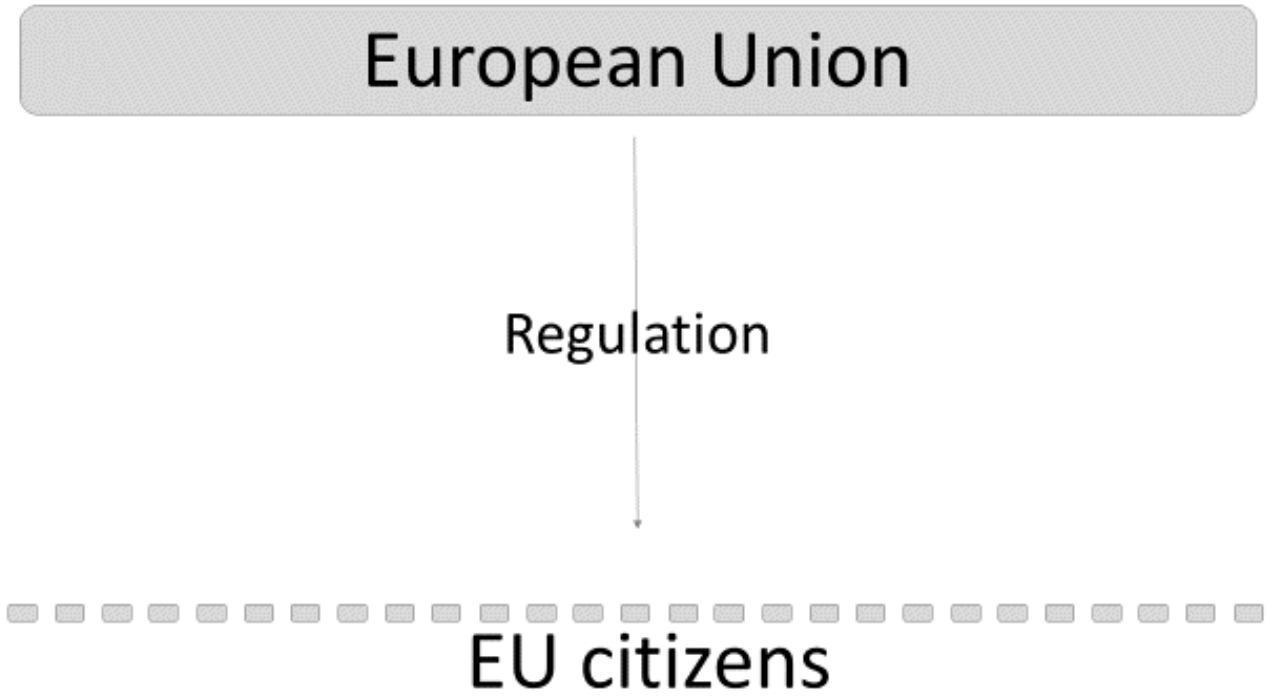

Figure 2.3: EU regulation

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, whether the processing takes place in the Union or not.

2. This Regulation applies to the processing of personal data of data subjects 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 such data subjects.

3. This 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."

Hence, many data processing operations by providers of services outside the European Union would thus fall into the scope of the European data protection law. The (proposed) recitals 19 and 20 highlight these intentions: [26]

> "(19) 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

[26] Cf. LIBE proposal, available at: http://www.janalbrecht.eu/fileadmin/material/Dokumente/DPR-Regulation-inofficial-consolidated-LIBE.pdf.

> through a branch or a subsidiary with a legal personality, is not the determining factor in this respect.
>
> (20) In order to ensure that individuals are not deprived of the protection to which they are entitled under this Regulation, the processing of personal data of data subjects residing in the Union by a controller not established in the Union should be subject to this Regulation where the processing activities are related to the offering of goods or services, irrespective of whether connected to a payment or not, to such data subjects, or to the monitoring of such data subjects. In order to determine whether such a controller is offering goods or services to such data subjects in the Union, it should be ascertained whether it is apparent that the controller is envisaging the offering of services to data subjects residing in one or more Member States 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. [27] 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, 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. [28] These goods do not need to be physical, but have a market value.
When the processing operation of the observation of the behavior of the person affected occurs, according to recital 21 the Article 3 Par. 2 (b) applies.

> "(21) In order to determine whether a processing activity can be considered to 'monitor' data subjects, it should be ascertained whether individuals are tracked, regardless of the origins of the data, or if other data about them is collected, including from public registers and announcements in the Union that are accessible from outside of the Union, including with the intention to use, or potential of subsequent use of data processing techniques which consist of applying a 'profile', particularly in order to take decisions concerning her or him or for analyzing or predicting her or his personal preferences, behaviors and attitudes"

For instance, when Internet activities are tracked by means of data processing techniques by which a person is assigned to a profile. Particularly affected are tracking-tools, which operate on the use of cookies [29] for example for Targeted Advertising. [30] Due to the altered wording of 'monitoring' in Article 3 par 2 (b), a selective observation is not covered.
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

[27] *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`

[28] ECJ, decision of 09/07/1992 - C-2/90, recital 26.

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

[30] *Peifer*, K&R 2011, 543 (543 ff.); *Rammos*, K&R 2011, 692 (692 ff.); *Klar*, ZD 2013, 109 (113).

international law according to Article 3 Par. 3. Pursuant to recital 22 this affects places such as diplomatic or consular missions. [31]
Hence, the former territorial principle of Article 4 of the Data protection Directive 95/46/EG has been abandoned in favor of a more market- and user-orientated model. [32] This very broad territorial scope of the proposal might cause a strong protection of European citizens' rights, since the offerer of services or goods is bound to European data protection law irrespective where he is established. For Cloud Computing this might lead 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 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 his 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 companies' customers). [33]
The affected person might assert his or hers rights because of the GDPRs broad claim of applicability more easily. [34] They do not have to worry about the location of the processors' servers anymore. [35]
However, this approach might go way 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. [36] Moreover, it is not probable that the EU could enforce data protection standards to providers seated outside the Union or without having business in the EU. European supervisory authorities are not able to act outside the Union. Art. 51 GDPR only states that:

> "1. Each supervisory authority shall be competent to perform the duties and to exercise the powers conferred on it in accordance with this regulation on the territory of its own Member State [...]"

There is no solution to this problem, so far. [37] Although Art. 25 GDPR states that a controller outside the Union that is affected by its data protection law shall designate a representative in the Union, there are no possibilities for sanctions or measures against such controllers in the GDPR; this has been criticized by the former German Federal Minister of Justice Sabine Leutheusser-Schnarrenberger. [38]

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.

[31] *Art. 29-Working Part*, Opinion 08/2010, WP 179, 22 ff.; *Wieczorek*, DuD 2013, 644 (648).
[32] *Härting*, BB 2012, 459 (462); *Piltz*, Datenschutzreform: aktueller Stand der Verhandlungen im Rat, 20/01/2014.
[33] *Hornung/Sädtler*, CR 2012, 638 (640).
[34] *Roßnagel/Richter/Nebel*, ZD 2013, 103 (104).
[35] *Nebel/Richter*, ZD 2012, 407 (410).
[36] *Spindler*, GRUR 2013, 996 (1003); *Spindler*, GRUR-Beilage 2014, 101 (107).
[37] *Hornung/Sädtler*, CR 2012, 638 (640).
[38] *Leutheusser-Schnarrenberger*, MMR 2012, 709 (710).

2.2.1 Personal data and encryption under 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 center 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, not being 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 acts or within the scope of the DPD. [39] In cloud computing, the *'personal data'* definitional issue is crucial with respect to *anonymized, pseudonymized* and *encrypted* data. Concerning encrypted data - be it encrypted while transmission, storage or computations - one of the issues refer to the problem if they still can be qualified as personal data.

2.2.1.1 Personal Data and Cloud Computing

As already outlined, the character of being personal data is crucial for the application of the Data Protection Directive. Hence, we have to take a closer look on the criteria for assessing this character:

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') [40]; 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. It depends on the circumstances in each individual case, if information can be qualified as 'personal data'. For instance, a common family name may not single someone out within a country but probably identifies 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*

[39] Directive 95/46/EC of the European Parliament and of the Council of 24/10/1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, available at: `http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTML`; German Federal Data Protection Act, available at: `http://www.gesetze-im-internet.de/englisch_bdsg/englisch_bdsg.html#p0061`.

[40] *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.

> *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 controller/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 is no longer qualified as personal data; otherwise, the data subject's consent is needed (see 3.11). In legal terms: Is the cloud-computing provider considered to be 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: [41]

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 debated, in particular if a so-called *absolute* or *relative* approach has to be the basis for assessing controller's abilities to identify a person.
In few words, the "absolute approach" assumes personal data if there is any chance for the data controller 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. [42] 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 necessary effort for the data controller as relevant in order to identify the data subject. [43] 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 [44] - or, at least, has reasonable chances of obtaining the decrypting key.
Despite its enormous practical impact this aspect has not been clarified yet, neither by the ECJ [45] nor the EC-Commission - even though the trend is beginning to favor a relative understanding

[41] *Hon/Millard/Walden*, The Problem of 'Personal Data' In Cloud Computing - What Information is Regulated?, The Cloud of Unknowing, Part 1, p. 13.
[42] *Art. 29-Working Party*, Opinion 04/2007, 7; OLG Hamburg, MMR 2008, 687 (688); *Pahlen-Brandt*, DuD 2008, 34 (38).
[43] *Dammann* in: Simitis, BDSG, par. 3, recital 32; *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 (377); *Hon/Millard/Walden*, The Problem of 'Personal Data' In Cloud Computing - What Information is Regulated?, The Cloud of Unknowing, Part 1, p. 46.
[44] *Spindler*, Verhandlungen des 69. Deutschen Juristentages, Band I, Gutachten, 2012, F 115, 116.
[45] *Kühling/Klar*, NJW 2013, 3611 (3614).

in the case law of some courts. [46] In contrary, some national authorities -for instance, the so-called *Düsseldorf Kreis* (a committee consisting of all German federal and regional supervisory authorities)- support the absolute approach [47], as well as some other authors. [48]
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: [49]

> *"'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. [50] One may note this instance while assessing British court decisions (such as the ones referred to above). Hence, the risk of inconformity with directive arises out of the provision in case the absolute approach prevails. Furthermore, it does not appear to be the 'usual' way in the EU-Member States to implement the DPD requirements of the term 'personal data' by expressly focusing on the controller's perspective only (which can be a sign for a relative understanding). [51] 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 derived from those single provisions.
A remarkable gradation was stated in the Austrian data protection law in par. 4 No. 1 DSG 2000: [52]

> *"'Data' ("Personal Data") [Daten" ("personenbezogene Daten")]: 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"*

[46] *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; *AG München*, ZUM-RD 2009, 413 (414) = BeckRS 2008, 23037;*OLG Hamburg*, MMR 2011, 281; *LG Wuppertal*, MMR 2011, 65 (66); *LG Berlin*, CR 2013, 471; different point of view *AG Berlin-Mitte*, ZUM 2008, 83 = K&R 2007, 600 (601); *VG Wiesbaden*, MMR 2009, 428 (432).

[47] `http://www.bfdi.bund.de/SharedDocs/Publikationen/Entschliessungssammlung/DuesseldorferKreis/Nov09Reichweitenmessung.pdf?__blob=publicationFile.`

[48] *Kuner*, European Data Protection Law, p. 92; *Marnau/Schlehahn*, Cloud-Computing: Legal Analysis, TClouds (D 1.2.2), p. 26 f.; *Pahlen-Brandt*, DuD 2008, 34 ff.

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

[50] 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.

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

[52] Austrian data protection act from 2000, BGBl. I Nr. 165/1999, last amendment 23/05/2013, BGBl. I Nr. 165/1999, english version available at: `https://www.dsb.gv.at/DocView.axd?CobId=41936.`

The Austrian law apparently 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). [53] 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. [54] This special category is called "indirectly personal" data by Austrian law. In other words: As long as identifiability has to 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 sub-category 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 Art. 2 (a) DPD). [55] 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. [56]
Regarding to 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, at all, 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. [57] As a consequence, even those data would be (directly) personal data and hence the data protection act would apply comprehensively without any privilege. In summary, the approach of the Austrian law should not be generalized, since it is on the one hand based upon a unique interpretation of the DPD assuming differences between a direct and an indirect identifiability and is on the other hand subject to a controversy about the actual scope of the term "indirectly personal" data.
The position of the Article 29 Data Protection Working Party [58] describes its stance concerning Art. 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*

[53] *Pollirer/Weiss/Knyrim*, Datenschutzgesetz 2000, par. 4, recital relating to Z 1, p. 20.
[54] *Pollirer/Weiss/Knyrim*, Datenschutzgesetz 2000, par. 4, recital relating to Z 1, p. 20f.
[55] Cf. *Bergauer*, Jahrbuch Datenschutzrecht 2011, 55 (60).
[56] Cf. *Bergauer*, Jahrbuch Datenschutzrecht 2011, 55 (57 f.).
[57] Cf. *Bergauer*, Jahrbuch Datenschutzrecht 2011, 55 (62).
[58] `http://ec.europa.eu/justice/data-protection/article-29/index_en.htm`.

> *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."* [59]

This opinion of the Working Party is interpreted by some authors as being cryptic as the Working Party used indirectly similar notions for other cases (mentioned in the opinion) which are pointing out more to the relative approach. [60] Others argue that the opinion includes a rather absolute stance. [61]
Indeed, if one takes a closer look at the Working Party's statement, it should be noted that it takes into consideration 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. [62] This instance can be regarded as an indication for an absolute approach. However, the Working Party apparently recognizes situations in which a set of data has to be regarded as *personal data* with respect to one entity but not with respect to another one [63], 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 (*whose* capacities have to be considered: Only the ones of the controller or of any other person in the world?). Hence, 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 [64], on the one hand - but, on the other hand, these means are limited to those which are reasonably likely to be used in the *concrete situation*. Only theoretical chances of identifying are insufficient to constitute the personal-characteristic of the data. [65]
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.
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 for assuming 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? [66] To put it simply, the 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.

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

The absolute approach (as a radical perspective) leads to a wide scope of Data Protection: only if there is no (theoretical) chance for the cloud-computing provider to re-combine the

[59] *Art. 29-Working Party*, Opinion 04/2007, WP 136, 21; see also Leonard, International Data Privacy Law, 2014, 53.
[60] Cf. criticism of *Kühling/Klar*, NJW 2013, 3611 (3614); *Pahlen-Brandt*, DuD 2008, 34 f.
[61] Cf. *Eckhardt*, CR 2011, 339 (341, 343); *Stimerling/Hartung*, CR 2012, 60 (63).
[62] *Art. 29-Working Party*, Opinion 04/2007, WP 136, 18 f.
[63] *Art. 29-Working Party*, Opinion 04/2007, WP 136, 15 f.
[64] Cf. also *Bygrave*, Data Privacy Law, p. 132.
[65] *Art. 29-Working Party*, Opinion 04/2007, WP 136, 15
[66] This distinction is made, for instance, in nuclear law and other laws referring to 'dangerous' technologies.

data in order to identify the data subject will the DPD not be applicable. In particular, encryption of data would not change the basic character of data (of personal data), itself, only render it quite difficult to access for unauthorized people. Hence, from a radical stance, encryption makes it more difficult to identify and "read" the personal data; however, it does not exclude the theoretical (!) chance 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 have (theoretically!) a chance to decrypt it. 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 people can be identified from the key-coded data when in combination with the key. If encryption were applied to a data set, the whole data set would be transformed and not just names within the data set. However, where the data controller possesses the decryption key, encrypted personal data might be viewed similar to key-coded data. If so, it would still be considered as 'personal data',.

2.2.1.4.2 The Impact of the Relative Approach upon 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. [67] 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 on the other side still processes personal data, since the assessment can vary depending on the particular person in question. [68]
In other words, the relative approach focuses on reasonable terms by which a provider may identify the data subject, particularly if it would be economically (and legally [69]) feasible. [70]
As possibilities/capacities of providers and their economic interest may vary widely, it cannot be assessed in general terms what qualifies as a reasonable effort to de-anonymize.
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

[67] *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.

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

[69] For instance, non-disclosure provisions or secrecy legislation may impede any re-combination of data between different providers. The supporter of the absolute approach negate these barriers.

[70] *Gola/Schomerus*, Bundesdatenschutzgesetz, par. 3, recital 44.

or de-anonymization is almost impossible. [71] If, under usual conditions, de-anonymization can be regarded as impossible, then the (anonymized) data cannot be qualified as "personal". [72]
Of course, the DPD will be applicable if the encryption still implies personal information; [73] we have to keep in mind that any personal identifier (even an IP-address) may be qualified as personal data.
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 even know if the information being processed while using their services is really 'personal data'. Hence, some authors argue that it might even seem inappropriate to apply the DPD to such cloud infrastructure providers as the processing of personal data depends upon their customers' choices. [74]
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. [75] Therefore, representatives of the relative approach [76] tend to apply the DPD to encrypted data as sooner or later, technical tools may facilitate decryption, like the encryption of DVDs. [77] Thus, foreseeable technical developments should be taken into account when assessing the current quality of personal data. [78] 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. [79]
However, 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 in order 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 technology. [80]

2.2.1.4.3 Conclusion

As the absolute approach extends the scope of DPD to nearly all kind of data processing [81] , from the perspective of the authors of this report (and from the perspective of most authors), the better arguments favour the relative approach. [82] Based on the absolute approach data

[71] *Gola/Schomerus*, Bundesdatenschutzgesetz, par. 3, recital 44; Kroschwald, ZD 2014, 75, (78).
[72] *Kühling/Klar*, NJW 2013, 3611, (3613); *Dammann*, in: Simitis, BDSG, par. 3, recital 32; *Gola/Schomerus*, Bundesdatenschutzgesetz, par. 3, recital 10; *Polenz* in: Kilian/Heusser, Computerrechts-Handbuch, Part 13, recital 68.
[73] *Spies*, MMR-Aktuell 2011, 313727.
[74] *Hon/Millard/Walden*, The Problem of 'Personal Data' in Cloud Computing - What Information is Regulated?, The Cloud of Unknowing, Part 1, p. 26.
[75] *Art. 29-Working Party*, Opinion 04/2007, WP 136, 15; *Kroschwald*, ZD 2014, 75 (78).
[76] *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.
[77] *Kroschwald*, ZD 2014, 75 (79).
[78] *Spies*, MMR-Aktuell 2011, 313727.
[79] *Stadler*, Datenschutz: IP-Adressen als personenbezogene Daten.
[80] *Kroschwald*, ZD 2014, 75 (78 f.); *Art. 29-Working Party*, Opinion 04/2007, WP 136, 15; Cf. also *Roßnagel/Scholz*, MMR 2000, 721 (723).
[81] *Meyerdierks*, MMR 2009, 8 (10); *Peifer*, K&R 2011, 543 (544); *Spindler*, Verhandlungen des 69. Deutschen Juristentages, Band I, Gutachten, 2012, F 115.
[82] *Dammann* in: Simits, BDSG, par. 3, recital 32; *Gola/Schomerus*, Bundesdatenschutzgesetz, par. 3, recital 10; *Spindler*, Verhandlungen des 69. Deutschen Juristentages, Band I, Gutachten, 2012, F 115, 116; *Schulz*

controllers (or data processors) cannot really assess if the DPD is applicable, since the DPD would be extended to an omni-present law without any real boundaries. [83] Furthermore, it should be considered that the purpose of the directive is in particular the protection of the right to privacy of natural persons (see Art. 2 No. 1 DPD). In scenarios where no realistic ("reasonable") chances to identify the data subject exist, with respect to the concrete situation of the processing actions, this purpose is not affected at all. Therefore, it does not seem necessary to apply restricting data protection under those circumstances. [84]
However, 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. [85] Even if at the beginning of data processing the data has not been 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. [86] Also, data which are related to things ("Internet of things") can turn out to be personal data if the data can be brought with reasonable effort [87] into a direct relationship with a person. [88]

2.2.2 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 labor, encryption technologies must be in a way sophisticated that efforts to attribute information to persons (to decrypt) must turn out unreasonable. 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 - not crucial is an objective point of view (if anyone in the world would be able to decrypt it).
Seeing strong support for the relative approach, even in the newly proposed General Data Protection Regulation, we have to underline that there is currently no judgment by a higher court (in particular by the ECJ) on this matter that confirms the relative approach.

2.2.3 Personal data and encryption under the GDPR

The definition of personal data has been greatly extended in contrast to the Directive 95/46/EG. [89] Article 4 Par. 2 now includes data, with which an indirect link can be made to a person: [90]

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.

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

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

[85] 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), 2012/0011 (COD) of 25/01/2012, recital 24: online identifiers combined with other information, available at: `http://www.ec.europa.eu/justice/data-protection/document/review2012/com_2012_11_eu.pdf`.

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

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

[88] *Art. 29-Working Party*, Opinion 04/2007, WP 136, 19 ff.

[89] The current definition of personal data in Article 2a of the Directive 95/46/EC is available at: `http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTML`.

[90] *Härting*, CR 2013, 715 (717).

> "(2) 'Data subject' means an identified natural person or a natural person who can be identified, directly or indirectly, by means reasonably likely to be used by the controller or by any other natural or legal person, in particular by reference to an identification number, location data, online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that person"

With the LIBE-proposal two new definitions have been added: The GDPR will provide precise definitions of 'pseudonymous data' and 'encrypted data' in Article 4 Par 2a and 2b.

> "(2a) 'pseudonymous data' means personal data that cannot be attributed to a specific data subject without the use of additional information, as long as such additional information is kept separately and subject to technical and organisational measures to ensure non-attribution;
>
> (2b) 'encrypted data' means personal data, which through technological protection measures is rendered unintelligible to any person who is not authorised to access it"

Unfortunately, the definition of 'encrypted data' does not exclude encrypted data from the applicability of the GDPR, in general, since the definition concerns 'personal data' that has been altered to be unintelligible. The direct effect encryption of data takes from a legal perspective, as intended by the GDPR, is relatively small. If data has been encrypted, the controller is not required to communicate a data breach to the data subject, according to Article 32 par. 3 GDPR. The notification requirements in Art 13 and 13a GDPR provide for an indication of whether or not the data processed will be encrypted. An indirect effect (not explicitly mentioned in the GDPR) that encryption might take on the processing of personal data could be the strengthening of the legitimate interests pursued by the controller during the balancing of interests required for an explicit legal permission to process data according to Article 6 Lit. f GDPR (see 2.4.2.2). 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, at all. This interpretation would support an absolute approach (see 2.2.1). However, it does not take into account that the qualification of data as personal or non-personal depends on the respective controllers' perspectives.
According to this approach (the relative approach, as described above, see 2.2.1), for the party able to decrypt the data, it still has to be considered personal data; whereas, for the party not able to decrypt it, the data is being considered anonymous. Hence, the norms of the GDPR that concern encrypted data are interpreted only as setting rules for the controller that is able to decrypt the data and how they should process it. In other terms, the norms do not mean that encrypted data always has to be considered personal data for every party. The GDPRs acknowledgment of encryption technologies and the benefit granted by Art 32 par. 3 to the controller who encrypts data can offer an incentive to controllers to encrypt the affected persons data before processing it. However, it does not answer the question of whether or not the encrypted data is considered personal data for a party that is unable to decrypt it. This still depends on the approach taken to define "identifiability" (see the following). Yet, the proposal seems to assume that the processing of encrypted data is less dangerous for the affected person's privacy than the processing of un-encrypted data (because the controller does not have to report a data breach to the data subject if the data was encrypted).

The GDPR will not be applicable to anonymous data: Recital 23 sentences 4, 5 clarify that the data protection legislation does not apply to anonymous data: [91]

> "Recital 23: [...] The principles of data protection should, therefore, not apply to anonymous data, which is information that does not relate to an identified or identifiable natural person. This Regulation does therefore not concern the processing of such anonymous data, including for statistical and research purposes."

Moreover, the anonymity is also mentioned in the context of health data in Article 81, so that they are not covered by the privacy regulation. Hence, an exact definition of when data become anonymized is not provided by the regulation rather than described by Recital 23. Unfortunately, this "definition" does not resolve the mentioned dispute between the different approaches (relative vs. absolute) to define anonymization. Therefore the same problems as mentioned persist such as upcoming techniques to decrypt or to identify data subjects by combining different pieces of information. [92] Techniques like removing or scrambling direct identifiers or even indirect identifiers, apparently cannot anonymize the data irreversible virtually. [93] Therefore according to the absolute approach almost all data have to be considered 'personal data'
It has been stated that, according to the GPDRs definition of 'personal data,' it is no longer important whether information relates only to a pseudonym which does not allow any conclusions about the real name. [94] For encrypted data this would mean that it is not important if the data has been encrypted or not: every piece of information which can be related to a person had, therefore, to be considered as personal data [95] - which would greatly extend the scope of the regulation on the European level. [96] 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 it has to be taken into account if the link between the affected person and the data can be made only with an extreme effort. This, too, is based on Recital 23 GDPR:

> "Recital 23: [...] To determine whether a person is identifiable, account should be taken **of all the means reasonably likely to be used** either by the controller or by any other person to identify or single out the individual directly or indirectly. [...]"

Hence, with a relative approach it can still be pointed out that the recital might take into account the means used by the respective controller **and** a third person - but only if those means are reasonably likely to be used. [97] If a decryption of the data is not reasonably likely to happen, the data could be considered non-personal (i.e. anonymous data) because the affected person would not be identifiable.
The LIBE version of the proposal has been provided with an explanation, written by draftsperson Jan Albrecht. It allows insight into the motives behind, at least, the LIBE version of the

[91] *Härting*, CR 2013, 715 (718).
[92] *Hon/Millard/Walden*, The Problem of 'Personal Data' In Cloud Computing - What Information is Regulated?, The Cloud of Unknowing, Part 1, p. 22.
[93] *Hon/Millard/Walden*, The Problem of 'Personal Data' In Cloud Computing - What Information is Regulated?, The Cloud of Unknowing, Part 1, p. 22.
[94] Specifying the problem of information relating to a pseudonym: *Härting*, Internetrecht, recital 185 ff.
[95] *Härting*, CR 2013, 715 (718).
[96] However, note that this extension depends on the former practice in member states. Germany already used a wider notion of personal data, even according to the so-called 'relative approach', see 2.2.1.
[97] *Lang*, K&R 2012, 145 (146).

GDPR. In this explanation, it is stated that the GDPRs purpose is to protect the fundamental rights of the affected persons. With that in mind, a limitation of the 'personal data' definitions scope is rejected. [98] All objective factors should be taken into account when determining if data is 'personal data,' according to the explanation. This is clearly a vote for an absolute approach, although it can be criticized for the same reasons as described above (see 2.2.1.4.3).

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 subcontractors whose resources they are using when their own capabilities are limited. [99] One has to 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, Art. 2 (d) DPD. It is not necessary for the controller, themselves , to process the data (see 2.3.1.4).
Two important elements included in this definition need to be described in more detail. First, the controller is the determining one - 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 analyzed. In case the role of the controller is assigned to one party or this one can be considered dominant relating to data issues altogether, this might be an important indication.

[98] *Albrecht*, Draft Report on the proposal for a regulation of the European Parliament and of the Council on the protection of individual 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 16/01/2013, 212, available at: `http://www.europarl.europa.eu/meetdocs/2009_2014/documents/libe/pr/924/924343/924343en.pdf`.

[99] *Brennscheidt*, Cloud Computing, p. 59.

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 concerned data subjects. [100]

2.3.1.3 Joint controlling

In a simplified data processing situation, there might only be one party held responsible, as a controller, when relating to the processing action. Nevertheless, the definition provided within Art. 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 the role of joint controllers. As a consequence, each of these parties are bound to the provisions stated within the DPD, with respect to the entire processing action. [101]
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). [102] 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. [103] However, in practice, the borderline between joint controlling, on the one hand, and order processing (see 2.3.1.4.2) of data, on the other hand, is hard to draw - and leads often 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. They can generally choose any legal form to establish their relationship - though, this does not affect the responsibility imposed by data protection law. [104] 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 needed in order to decide the issue if parties take the decisions jointly, or if only one party has to be regarded a ("single") controller. [105]
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. [106]
The legal assessment is unambiguous, 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. [107]
Furthermore, as the Art. 29 Working Party argues, the question of joint controlling is not a matter of one particular data processing action. As Art. 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). [108]
Especially in the context of IT-infrastructures, there can be many parties involved in different

[100] *Art. 29-Working Party*, Opinion 01/2010, WP 169, 8 ff.
[101] *Wolff/Brink*, Datenschutz in Bund und Ländern, par. 3, recital 112.
[102] *Art. 29-Working Party*, Opinion 01/2010, WP 169, 18.
[103] *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).
[104] *Dammann*, in: Simitis, BDSG, par. 3, recital 226.
[105] *Art. 29-Working Party*, Opinion 01/2010, WP 169, 18; see also 2.3.1.2.
[106] *Jandt/Roßnagel*, ZD 2011, 160, Jotzo, MMR 2009, 232 f.
[107] *Art. 29-Working Party*, Opinion 01/2010, WP 169, 19 f.
[108] *Art. 29-Working Party*, Opinion 01/2010, WP 169, 18.

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 a way that each single action appears to be independent and executed by only one controller. However, the entities can also be regarded as joint controllers by taking into consideration the whole set of operations - the "macro-level". This result can be derived from jointly determined purposes, from a jointly set framework that determines the essential means or if the decisions relating to both questions are taken together. [109] 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 the issue: [110] An airline, a hotel chain and a travel agency establish a platform provided through the internet that allows a better 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 keep in mind, that the Art. 29-Working Party opinions have no binding statements (see Art. 29 section 1 DPD). In particular, it may be subject to further discussion if such a wide understanding of joint controlling can generally be accepted. The ECJs 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. [111] Simply the fact that both parties were able to control the processing had been sufficient for the ECJ to assume joint controllership. [112]
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 (Art 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".

[109] *Art. 29-Working Party*, Opinion 01/2010, WP 169, 20.
[110] *Art. 29-Working Party*, Opinion 01/2010, WP 169, 20.
[111] ECJ, Judgment from the 13th May 2014 in Casec-131/12 - Google Spain SL/Google Inc. v AEPD/Mario Costeja Gonzalez, rec. 40
[112] Cf.*Spindler*, JZ 2014, coming soon.

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 Art. 16 DPD), at least with respect to the purposes of the processing and the essential means that are used. In this respect, 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". [113]
Furthermore, only an entity legally separated from the controller is in general able to act as a processor. [114]

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 between these two roles has to been 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). [115] So, if one determines the purposes and essential means (at least by giving instructions) he is a controller. [116] In this context, it is crucial to specify which particular decisions can be delegated to the processor, in contrast how much leeway or discretion is assigned to the processing party so that it can be already 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 can be 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. [117] As a consequence, the cloud service provider will be considered a controller, themselves, if they collect their users' personal data for their own purposes. [118]
Decisions that concern the means of the processing might, on the other side, be delegated to the processor, in general; for example, which software should be used. However, this does not include 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 relates especially to the duration of the processing, granting access to third persons, and the choice of which data should be processed. [119]
In a typical cloud computing scenario, the provider only provides the technical framework which is used by the controller. The latter is the one determining the purposes of the processing. Usually, this one decides which data are processed and how long the processing will take and, therefore, governs the (essential) means, whereas the cloud provider only computes the data, as they are bound by the contract concluded with the cloud user, thus having little discretionary power that, normally, does not lead to a controllership. [120]

[113] *Art. 29-Working Party*, Opinion 01/2010, WP 169, 13 f., 25.
[114] *Art. 29-Working Party*, Opinion 01/2010, WP 169, 25.
[115] *Brennscheidt*, Cloud Computing und Datenschutz, p. 67; Gola/Schomerus, Bundesdatenschutzgesetz, par. 11, recital 9.
[116] Cf. *Hilber*, Handbuch Cloud Computing, p. 350
[117] *Art. 29-Working Party*, Opinion 01/2010, WP 169, 15 f.
[118] *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.
[119] *Art. 29-Working Party*, Opinion 01/2010, WP 169, 14.
[120] *Brennscheidt*, Cloud Computing und Datenschutz, 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

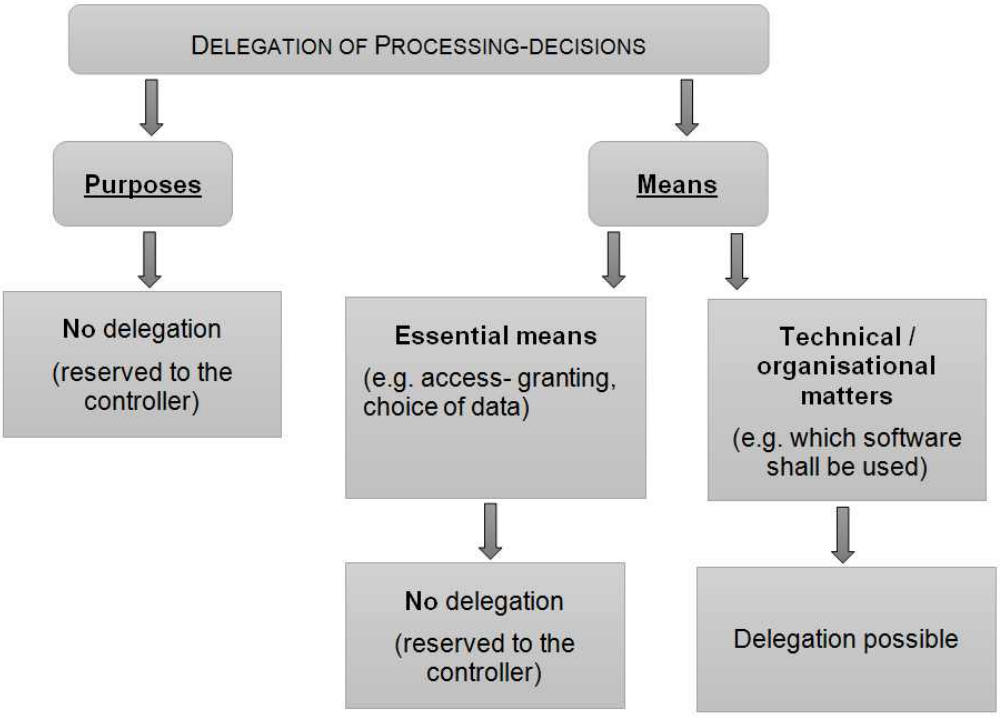

Figure 2.4: Order-processing - Delegation of decision

Even though cloud computing can, therefore, usually be regarded as processing on behalf of the controller, in terms of Art. 16 DPD [121], it is discussed 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 an 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. [122] 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. [123] 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. [124]
However, this discussion might not be overrated. It should be kept 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.

Rechtsfragen des Cloud Computing, chapter D, recital 31 ff.

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

[122] *Hon/Millard/Walden*, Who is Responsible for 'Personal Data' in Cloud Computing?, The Cloud of Unknowing, Part 2, p. 17; *Spindler/Schuster*, Recht der elektronischen Medien, par. 11 BDSG, recital 7.

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

[124] *Hon/Millard/Walden*, Who is Responsible for 'Personal Data' in Cloud Computing?, The Cloud of Unknowing, Part 2, p. 22.

Accordingly, the provider has to be considered a processor. [125] This applies all the more if the provider fulfills monitoring tasks with respect to the personal data, e.g. concerning the access or use. [126]
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. [127] The same might apply if they exceed other competences, such as granting data access to unauthorized third parties. [128] Furthermore, the provider could be assigned not only with providing the technical framework but also with completing the complete task that leads to the processing action. Whenever the 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. [129] The outsourcing of a company's accountancy is a typical example, in this respect. [130]

2.3.1.4.3 Legal requirements

There are certain legal requirements to fulfill before (order) processing takes place 'on behalf of the controller' (Art. 17 Par. 3 DPD), such as 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 oblige the processor to follow the controller's instructions, similar to an employee's obligation. For the purposes of keeping proof, the parts of the contract or the legal act relating to data protection and the requirements relating to the technical and organizational measures shall be in writing or in another equivalent form. [131] 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 is still 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. [132]
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

[125] *Pohle/Ammann*, K&R 2009, 625 (630); *Spindler/Schuster*, Recht der elektronischen Medien, par. 11 BDSG, recital 7; 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.
[126] *Hon/Millard/Walden*, Who is Responsible for 'Personal Data' in Cloud Computing?, The Cloud of Unknowing, Part 2, p. 17.
[127] *Art. 29-Working Party*, Opinion 05/2012, WP 196, 14.
[128] *Hon/Millard/Walden*, Who is Responsible for 'Personal Data' in Cloud Computing?, The Cloud of Unknowing, Part 2, p. 20.
[129] Cf. *Brennscheidt*, Cloud Computing und Datenschutz, p. 67;*Funke/Wittmann*, ZD 2013, 221 (223); Hoeren, in: *Roßnagel*, Handbuch Datenschutzrecht, chapter 4.6, recital 97.
[130] *Petri*, in: Simitis, BDSG, par. 11, recital 28.
[131] *Art. 29-Working Party*, Opinion 05/2012, WP 196, 12.
[132] *Art. 29-Working Party*, Opinion 01/2010, WP 169, 26; *Hilber*, Handbuch Cloud Computing, p. 357.

- 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 [133]

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 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. [134] 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 to be impossible 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. [135] 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 fulfil 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). [136] On the other hand, the cloud user would still have to pay for 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. [137] An effective, yet practical, way to ensure compliance is data protection certification. [138] Here, as well, 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, and affordable even 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

[133] The whole list of recommendations has 14 items and can be found in Art. 29-Working Party, Opinion 05/2012, WP 196, 12 f.
[134] *Heidrich/Wegener*, MMR 2010,803 (806).
[135] *Brennscheidt*, Cloud Computing und Datenschutz, p. 102.
[136] *German Federal Office for Information Security Technology*, Safety Recommendation for Cloud Computing Providers, p. 63.
[137] *Brennscheidt*, Cloud Computing und Datenschutz, p. 105
[138] *Art. 29-Working Party*, Opinion 05/2012, WP 196, 22.

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. [139]

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 (see 2.4.3) requirements. Nevertheless, the contract binding the processor to the controller can be used to ensure necessary safeguards. So, in other words,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. [140]

2.3.2 The responsible party (the controller) and processing on behalf of the controller under the GDPR

The GDPR, too, distinguishes between the entity responsible and the entity actually processing the data. Nevertheless, there will be changes in the particular responsibilities of those entities and new ways for the controller to make sure his processor complies with the law. Order processing under the GDPR has to meet all prerequisites described below.
It has been criticized that there is no regulation within the GDPR explicitly stating that transfers from a controller to his processor are allowed, in general, if 'order processing' takes place. [141] 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 senseless. [142]
In general, the processing of data by the processor (like 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 et seq.).

2.3.2.1 Rules for the controller - Article 22

The controller is defined in Art. 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; [...]".

[139] For a detailed description of data protection seals see *Brennscheidt*, Cloud Computing und Datenschutz, p. 105 ff.
[140] *Brennscheidt*, Cloud Computing und Datenschutz, p. 76.
[141] *Nebel/Richter*, ZD 2012, 407 (411);*Roßnagel/Nebel/Richter*, ZD 2013, 103 (105); c.f. *Koós/Englisch*, ZD 2014, 276 (284), who see the legitimation in Art 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.
[142] C.f. regarding the DPD, but with the same problem: *Drews/Montreal*, PinG 14, 143.

There will be no significant changes in the definition of 'controller,' compared to the DPD 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) thus is responsible for the data processing and will be accountable if legal requirements are not met. The controller's main duties are regulated in Art. 22 of the GDPR.

> "Article 22 Par 1: The controller shall adopt appropriate policies and implement appropriate and demonstrable technical and organizational measures to ensure and be able to demonstrate in a transparent manner that the processing of personal data is performed in compliance with this regulation, i.e. having regard to the state of the art, the nature of personal data processing, the context, scope and purposes of the processing, the risks for the rights and freedoms of the data subjects and the type of the organization, both at the time of the determination of the means for processing and at the time of the processing itself"

Simplified, this means that it is the controller's (the cloud user's) job to ensure that the GDPRs 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.
Specifically, the criteria of the 'state of the art' and 'the risks for the rights and freedoms of the data subjects' can be addressed efficiently by using privacy preserving cloud computing technologies developed by PRACTICE.
Besides the obligation to ensure compliance and to provide policies that respect the data subjects' free choices (Art 22 Par 1a), the controller also has to be able to demonstrate the adequacy and effectiveness of those measures and policies. To achieve this, Recital 60 GDPR recommends independent internal or external auditors (see 2.3.2.4). Article 28 requires documentation of the data processing by the controller (and the processor, as well). They must cooperate with the supervisory authority of Article 29; take technical and organizational measures to ensure the security of processing Article 30; alert and inform clients about data breach, according to Article 31 Par. 2; conduct a privacy impact assessment under certain conditions of Articles 32a, 33 Par. 1 or seek a prior authorization in accordance with Article 34 Par. 1; appoint a data protection officer, as requested in Article 35 Par. 1; as well as comply with rules for transfers to third countries, as mentioned in Article 40 ff. The powers of regulators may be expressly addressed to the processors, according to Article 53 par 1 (a).

2.3.2.2 Joint Controllers - Article 24

The GDPRs definition of 'controller' allows several entities to be considered as 'joint controllers'. Since there have been only slight changes in the GDPRs definition of 'controller,' compared 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 such a 'joint controllers' scenario, it might be difficult to determine the certain 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 the arrangement has to determine which controller is responsible for the procedures and mechanisms for 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. Also, one controller could have a direct relationship to the data subject, whereas another one might not; as well as. they might not control the kind and amount of data. [143] The arrangement the GDPR demands should be seen as a useful tool for cloud participants when they are considered joint controllers. It might be included in the contract, determining the cloud service's details between the cloud provider and the cloud user(s). If the respective responsibilities are not clear to the data subject, all joint controllers are liable, together or separately. In this specific case, this is meant to provide the data subject with more protection. [144]

2.3.2.3 Rules regarding the processor - Article 26

In the relationship of the cloud provider and the cloud user, the cloud provider usually acts as the processor being 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; [145] we have to keep in mind that the cloud user often offers services to their clients, thus has to be qualified as a data processor (Client - cloud user - cloud provider).
The controller's duties regarding the processor begin before the processing on their behalf takes place: they have to choose a processor who will comply with the GDPRs requirements:

> "Art. 26 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, in particular in respect of the technical security measures and organizational measures governing the processing to be carried out and shall ensure compliance with those measures."

Art. 30 GDPR clarifies what is meant by 'technical and organisational measures' (see 2.4.4.2). Those measures shall, among other things, at least "protect stored or transmitted personal data against accidental or unlawful destruction, or accidental loss or alteration, and unauthorised or unlawful storage, processing, access or disclosure" (Art. 30 par. 2 lit. b GDPR). This is possible with encryption technology.
If the cloud provider uses privacy-preserving technologies, as they are developed by PRACTICE, it can be assumed the provider thus partially fulfils their duties regarding aforementioned technical security measures.

[143] *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`.

[144] *Comi*, IMCO Committee 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`.

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

Of course, they still have to take organizational measures to fulfil all of their duties regulated in Art. 30 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 make sure that they ensure their 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; Art. 26 (2) defines a set of rules that must, in practice, be endorsed in the contract, such as:

(a) process personal data only on instructions from the controller, unless otherwise required by Union law or Member State law;

(b) employ only staff who have committed themselves to confidentiality or are under a statutory obligation of confidentiality;

(c) take all required measures pursuant to Article 30;

(d) determine the conditions for enlisting another processor only with the prior permission of the controller, unless otherwise determined.

(e) insofar as this is possible, given the nature of the processing, create in agreement with the controller the appropriate and relevant technical and organizational requirements 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) assist the controller in ensuring compliance with the obligations, pursuant to Articles 30 to 34, taking into account the nature of processing and the information available to the processor;

(g) return all results to the controller after the end of the processing, not process the personal data otherwise, and delete existing copies unless Union or Member State law requires storage of the data;

(h) make available to the controller all information necessary to demonstrate compliance with the obligations laid down in this Article and allow on-site inspections;"

Although the data protection law will be renewed, the practical problems will still be the same. The cloud user, as 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 might 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, Art. 26 (4) 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, and shall be subject to the rules on joint controllers laid down in Article 24."

The switch of roles for the processor (from mere processing to determining and controlling any data processing) leads thus 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 Art 26 (2 d) that the processor may use services of other processors if the data controller has given their prior consent. [146] 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 have to assure that his inspection rights, etc., are also enforceable in the relationship with the third-party processor (sub cloud provider).

2.3.2.4 Privacy Seal - Article 39 GDPR

According to Article 39, the data protection authority can act as a certification authority. Each controller and data processor has the right to apply for a certification procedure as mentioned in Article 39 a, b.

> "(1a) Any controller or processor may request any supervisory authority in the Union for a reasonable fee, taking into account the administrative costs, to certify that the processing of personal data is performed in compliance with this Regulation– in particular with the principles set out in Article 5, 23 and 30, the obligations of the controller and the processor, and the data subject's rights,
>
> (1b) The certification shall be voluntary, affordable, and available via a process that is transparent and not unduly burdensome."

The certification procedure, however, may turn out to be, in practice, one of the most important tools for data controllers to bring evidence required by Article 26 (1), concerning the selection of processors with sufficient guarantees for data protection, particularly appropriate technical and organizational measures. [147] This might partly be a solution for the dilemma arising from the disparity of power between the cloud computing participants: the cloud provider will be able to request a certification the cloud user is allowed to rely on. However, there is no obligation for certification. [148] Moreover, Article 39 (1d) provides for third party certification procedures if the data protection authority has accredited them. [149]

> "(1d) During the certification procedure, the supervisory authority may accredit specialised third-party auditors to carry out the auditing of the controller or the processor on their behalf. [...]"

A certificate of the processor (issued by an accredited third party) may thus be considered as evidence in order to prove the compliance with these obligations. Not only the processor can request a certification; the controller might have an interest in getting certified, too. A cloud user (as the controller) might be able to prove to his clients that he uses a cloud service that is compliant with data protection law and that, especially, provides sufficient technical and organisational safeguards. The Commission will be empowered to adopt delegated acts (see 2.5.2) to further specify the criteria and requirements for the certification mechanisms,

[146] Brennscheidt, Cloud Computing und Datenschutz, p. 116.
[147] *Brennscheidt*, Cloud Computing, p. 116
[148] *Härting*, CR 2013, 715 (720).
[149] *Brennscheidt*, Cloud Computing und Datenschutz, p. 116; *Härting*, CR 2013, 715 (720).

according to Art 39 Par. 3 GDPR. Although the certification of the processor can make it much easier for the controller to bring evidence required by Article 23 (1), a certificate will expire after five years.

> (1g) Notwithstanding paragraph 1f, the certification shall be valid for maximum five years.

Before relying on a certificate, the processor will, therefore, at least be obliged to validate if it is expired or not. [150] Nevertheless, the problem arising from the fact that the cloud user has to ensure by contract that he controls the provider if an 'order processing' shall take place is not solved by a certification. [151] A cloud provider (even if he is a big global player) should make it easier for his client - the cloud user - to fulfill his obligations by providing standard contracts for his cloud services that involve the requirements of Art. 26 (2). This way, a lawful use of the cloud service would be possible for the cloud user, as a controller, if he wants to compute personal data in the cloud. If a lawful usage of a cloud service for the computation of personal data is not possible due to the cloud provider denying to contract with his clients (the cloud users, controllers) that, in a way, binds himself compliant to Art. 26 (2) out of his more powerful position, this might lead to a disadvantage on the European market once the GDPR comes into effect.

2.3.2.5 Liability - Article 77

If data has been processed unlawfully, the data subject has the right to claim compensation, even for non-pecuniary damages, according to Art. 77 GDPR. 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 77 (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 might 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 includes the possibility to avoid liability for damages.

> "(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 who needs to prove that he is not responsible for the damage, they both should take the technical and organizational measures that the GDPR demands and fulfill their duty to document the processing, according to Art. 28. It is an advantage of the affected person claiming compensation for damages that it is up to the processing parties to prove that they are not responsible. 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

[150] *Sydow/Kring*, ZD 2014, 271 (275).
[151] *Sydow/Kring*, ZD 2014, 271 (275).

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 processor's internal procedures. [152] Another advantage that will be provided to the affected person if several processors or joint controllers caused the damages is as follows:

> "(2) Where more than one controller or processor is involved in the processing, each of those controllers or processors shall be jointly and severally liable for the entire amount of the damage, unless they have an appropriate written agreement determining the responsibilities pursuant to Article 24"

The possibility of joint liability for joint controllers makes it important for them to come to an agreement that fully reflects their responsibilities in the data processing. This way, only the respective controller might be liable for damages caused by his actions.

2.3.2.6 Commissioned Data Processing in Third Countries - Article 3 Par. 1, 2

Cloud computing service, even those based outside Europe, may become subject to the European Data Protection Regulation (see 2.1.2.2). The scope of the proposed Regulation is not restricted to companies that have an establishment in the EU, according to Article 3 Par. 1. [153] Article 3 Par. 2 even extends the scope of the regulation to providers outside of the European Union which process the data of European Citizens (Article 3 Par. 2 (a) and (b)). [154] Even pure targeting and gathering data of EU-citizens by companies outside the EU is now covered by Article 3 Par 2 (a). [155]
As further outlined in 2.4.3.2, a data transfer to a processor in a third country has to be evaluated in two steps according to the DPD. First, a data transfer needs permission, which can be provided if the 'order processing' meets all requirements described above. Second, the transfer into the third country has to be legal, which will be further elaborated in 2.4.3.2.

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).

The extremely broad definition of 'processing' leads to the applicability of the DPD and, thus, to the general prohibition of processing the data unless the DPD allows for it. From the moment

[152] *Roßnagel/Richter/Nebel*, ZD 2013, 103 (108).
[153] *Härting*, BB 2012, 459 (462); *Wieczorek*, DuD 2013, 644 (648).
[154] *Hon/Millard/Walden*, The Problem 'Personal Data' in Cloud Computing - What Information is Regulated?, The Cloud of Unknowing, Part 1, p. 3; *Klar*, ZD 2013, 109 (112).
[155] *Härting*, BB 2012, 459 (462); *Wieczorek*, DuD 2013, 644 (648); *Klar*, ZD 2013, 109 (112).

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 making the data "not personal". Otherwise, they can comply with the requirements - asking the user for an explicit consent or bringing forward reasons that fall under the justifications provided by the DPD.
If personal data is anonymized, this might, technically, mean that it gets altered, but, for the purposes of the Data Protection Directive, 'alteration' means changing the information's content, not its appearance. [156]
Since the anonymization of personal data eliminates the connection to a person, the encryption of data is one of the few possibilities to anonymize personal data and, therefore, the process of encrypting data does not fall under the data protection law, either.

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 Art. 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 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 they are not just examples among other possible legal grounds but the only lawful reasons 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

[156] See also *Gola/Schomerus*, Bundesdatenschutzgesetz, par. 3, recitals 30, 31.

principle of proportionality laid down in Art. 6 DPD (see 2.1.1.2.2.3). Whereas lit (b) to (e) are applicable only for specific purposes, lit (f) allows the member states to provide a legal ground with a larger scope. But, minding Art. 6 DPDs main principles, a processing on the grounds of a permission based on lit (f) always requires a proportionality test. This means a balance has to be found 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 might be useful if data could be processed without consent. Cloud Computing frees the cloud user from providing his own physical resources needed for the data processing he wants to do. Instead he is able to use scalable cloud resources on demand when making use of a cloud provider's infrastructure or even software. This might hold financial advantages for the cloud user. It is questionable, though, if financial advantages are sufficient to outweigh the data subject's interest in a comprehensive data protection. [157] This interest is based on a European fundamental right, Art. 7 and 8 CFR. [158]
In a recent decision, the ECJ had to evaluate a similar conflict of interests: a spanish citizen demanded that Google be required to remove personal data concerning the search results and stop making the relevant search results available to the public. [159] He argued that Google no longer could base the processing of the data on the legal grounds of Art. 7 lit (f). The ECJ stressed the strong position of the data subject in such a balance of interests, stating that, in this case, merely the economic interest of Google would not be able to justify the processing. [160] The court hold that the rights of the data subject override, as a rule, the economic interests of the operator of the search engine. [161] Although the decision was not concerned with a balance of interests between a cloud user and an affected data subject, ratther than with a data subject who was facing a disadvantageous Google search result (linking information to him), the ECJs reasoning has to be minded for cloud computing, as well. Therefore, a processing on the grounds of Art. 7 lit (f) should not only be justified by a financial advantage of the cloud user.

2.4.2.2 Legal permissions in the GDPR

The proposal for a GDPR includes permissions for a processing of personal data that work 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), Art. 6 GDPR mentions five other exceptions:

> "[. . .]

[157] 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.

[158] 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`.

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

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

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

(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 to whom the data is disclosed, and which meet the reasonable expectations of the data subject, based on his or her relationship with the controller, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject, which require protection of personal data. This shall not apply to processing carried out by public authorities in the performance of their tasks.

[. . .]"

Actually, the DPD allows for data processing only if the affected party consents to it or if the data processing is necessary for legitimate purposes of 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. [162] 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 overridden by the data subjects interests; thus, data processing in order to enforce a statutory claim might be lawful without consent of the affected person under certain circumstances. Although, if the processing is permitted according to lit. (f), the data subject is able to object to the processing at any time, and without any further justification, free of charge, Art 19 Par. 2 GDPR. This broad right to object does not exist if the processing is based on lit (b), which might cause lit (b) to be the more reliable reasons for processing.

2.4.2.3 Informed Consent and Cloud Computing

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 Art. 7, consent is the first out of seven legal grounds for personal data processing.
According to Art. 8 of the directive, consent needs to be given explicitly for processing special categories of data. Some Member States see consent as a preferred ground for lawfulness, whereas others see 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 even beyond what is necessary for his purposes; [163] in other terms, the data processor is not bound by a strict proportionality test under these circumstances. [164] However, as noted already, the

[162] *Berg*, PinG 2013, 69 (70).
[163] *Art. 29-Working Party*, Opinion 15/2011, WP 187, 7f
[164] *Nägele/Jacobs*, ZUM 2010, 281 (290); *Rath/Rothe*, K&R 2013, 623 (624).

DPD treats specific sensitive data in a 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 given freely, informed 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 (Art. 10 (c) Data Protection Directive). [165] It also has to be made clear to the data subject when data will be transferred to a non-EU-state. [166]
These requirements of ex-ante information and transparency lead to difficulties: 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. [167] 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. [168] 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. consent. If consent is given by accepting terms and conditions, legal requirements regarding consumer protection law have to be met, as well. [169]
To ensure compliance with the data protection law, consent would not be the best solution in such a case. [170] 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. [171] This way, the data subject can be informed specifically about the location of his data. The same principle applies to the provider's sub-contractors: the user should choose which subcontractor he will use for the specific computation, so informed consent is ensured. Nevertheless, this might be impossible for cloud computing services using resources of other cloud providers. For instance, SaaS provider, Dropbox, builds his 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). [172] In such cases, cloud users will not necessarily know in which data

[165] *Taeger*, in Taeger/Gabel, BDSG, par. 4a, recital 30; *Nord/Manzel*, NJW 2010, 3756 (3757).
[166] *Simits*, in Simits, Bundesdatenschutzgesetz, par. 4a, recitals 70 ff.
[167] *Nägele/Jacobs*, ZUM 2010, 281; *Schultze-Melling*, in Taeger/Gabel, BDSG, par. 9, recital 104.
[168] *Millard* , Cloud Computing, Chapter 1.1, 1.2; *Funke/Wittmann*, ZD 2013, 221 (222).
[169] *Spindler*, GRUR-Beilage 2014, 101.
[170] BITKOM, Leitfaden Cloud Computing, p. 51; German Federal Office for Information Security Technology, Safety Recommendation for Cloud Computing Providers, p. 73; Brennscheidt, Cloud Computing und Datenschutz, p. 152; Art. 29-Working Party, Opinion 15/2011, WP 187, 12.
[171] For an example such a service is provided by Amazon Web Services, available at: `http://aws.amazon.com/de/ec2/pricing/effective-april-2014/`
[172] *Millard*, Cloud Computing, Chapter 3.2.

centers, or even countries, their data is stored or with whom their provider has a subcontractor relationship. Actually, the provider, himself might not know this. [173]

2.4.2.4 Informed Consent and obligation of transparency under the GDPR, Article 14

Art. 14 extends the approach of the DPD concerning transparency for data subjects (and also goes beyond existing national laws, like in Germany) [174] by specifying to the data subject the information that has to be handed prior to the collection of data. For the purposes of cloud computing, these obligations to inform raises 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:

> "(b) the purposes of the processing for which the personal data are intended, as well as information regarding the security of the processing of personal data, including the contract terms and general conditions where the processing is based on point (b) of Article 6(1) and, where applicable, information on how they implement and meet the requirements of point (f) of Article 6(1);
>
> (c) the period for which the personal data will be stored or, if this is not possible, the criteria used to determine this period;
>
> (d) the existence of the right to request from the controller access to and rectification or erasure of the personal data concerning the data subject, to object to the processing of such personal data, or to obtain data;
>
> (f) the recipients or categories of recipients of the personal data;
>
> (g) where applicable, that the controller intends to transfer the data to a third country or international organization and on the existence or absence of an adequacy decision by the Commission, or in case of transfers referred to in Article 42, Article 43, or point (h) of Article 44(1), reference to the appropriate safeguards and the means to obtain a copy of them;
>
> (ga) meaningful information about the logic involved in any automated processing;
>
> (h) any further information which is necessary to guarantee fair processing in respect of the data subject, having regard to the specific circumstances in which the personal data are collected or processed, in particular the existence of certain processing activities and operations for which a personal data impact assessment has indicated that there may be a high risk;
>
> (ha) where applicable, information whether personal data was provided to public authorities during the last consecutive 12-month period."

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 really 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 - what 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

[173] 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`

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

the data, it will be even more important for him that the data is not considered 'personal data' under the GDPR.
The controller has to provide the information, according to Article 14 GDPR, before personal data is collected - which also includes the collecting of data based on an explicit legal permission, as provided by Article 6 (b) - (f). In contrast, the DPD just requires the controller to provide such information to gain informed consent. The GDPR, on the other hand, requires that such information is also provided before data is collected, on the grounds of a legal permission.
The cloud user (controller) can comply with these obligations by choosing a cloud provider who enables him to determine which servers in what country will be used to offer the cloud service. [175] Whereas the DPD only requires a freely given consent, particularly an informed consent, the GDPR demands much more from a controller.
These obligations to inform are flanked by the new provision in Art. 13a of the Regulation which requires the data controller to provide standardized and easily legible information (what is, in detail, prescribed by the annex of the proposed Regulation). Concerning the information of Art. 14, 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.

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. [176] 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. [177] Hence, there is a difference between the legal permission to process data and the legal permission to transfer the data into a third country. Only if both requirements are fulfilled separately is the data transfer lawful.
One of the main exceptions refers to the 'adequate level of protection' in the third country, Art. 25 of the Directive. [178] An adequate level of protection assumes that the data protection standards in the respective country are comparable to the European standards. This has to be officially acknowledged by the European Commission as for instance for: Andorra, Argentina, Australia, Canada, Switzerland, Faeroe Islands, Guernsey, the State of Israel, the Isle of Man, Jersey, New Zealand, the United States - Safe Harbor, and Eastern Republic of Uruguay. [179] Being of particular relevance for cloud computing, the USA has not been acknowledged, in general. [180]

[175] 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/`.
[176] 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.
[177] See also *Brennscheidt*, Cloud Computing und Datenschutz, p. 175.
[178] *Hon/Millard*, Data Export in Cloud Computing - How can Personal Data be Transferred outside the EEA?, p. 5.
[179] 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`
[180] No such decision has been made by the commission; *Gabel*, in: Taeger/Gabel, BDSG, par. 4b, recital 23; BITKOM, Leitfaden Cloud Computing, p. 53.

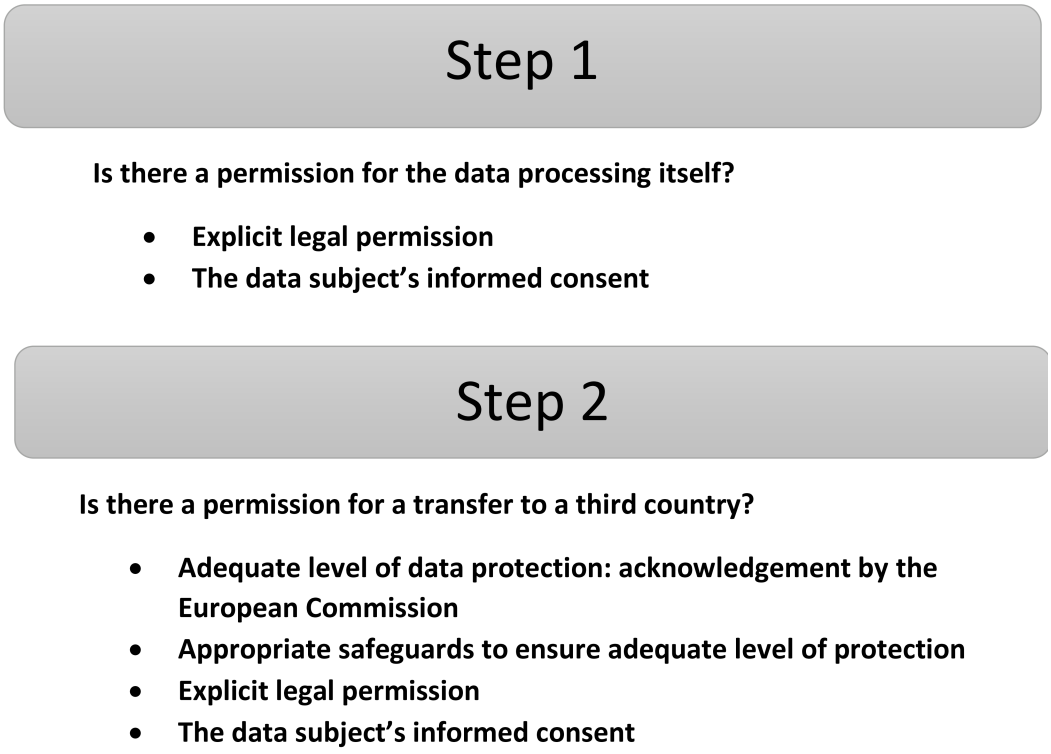

Figure 2.5: Data transfers to third countries

Besides these officially acknowledged countries (where no explicit consent by the user is needed), the data controller who wants to transfer data in other countries may use other forms of justification provided by the DPD, this is possible, in general, if the controller adduces evidence of adequate safeguards with respect to the protection of the data subject's rights, Art. 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 [181] 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 are equally protecting the data subject's rights as the EU data protection law does. Yet, the benefit of a lawful transfer to the third country only exists if the clauses acknowledged by the commission are used exactly like the commission provided them and without alteration. [182] For a cloud user who wishes to transfer data to a cloud provider within a third country, the standard contractual clauses would then only be useful if the cloud provider agrees to those exact clauses. It seems unlikely that a cloud provider contracting with many cloud users would alter his normal contractual agreements, but would, instead, agree to the standard contractual

[181] 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`.

[182] *Gola/Schomerus*, Bundesdatenschutzgesetz, par. 4c, recital 14; *Spindler/Schuster*, Recht der elektronischen Medien, par. 4c BDSG, recital 20.

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, Art. 26 par. 2 is not exhaustive, which means appropriate safeguards might be other measures than the explicitly mentioned standard contractual clauses; they are only an example among other possible safeguards. [183] 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." [184]

The BCR have to be binding or legally enforceable and should be regarded as "sufficient safeguards" within the meaning of Art. 26 par. 2 DPD. They are meant to be used by multinational companies to allow international data transfers. [185] There are no model BCR provided by the Art. 29-Working Group or the commission, like 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. [186] 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). [187] In some specific cases, BCR may be used for cloud computing-related data transfers, however, the will be restricted to internal data transfers across borders. [188] 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. [189]
Although the Safe Harbor agreement has recently been criticized for not living up to the requirements of European Data Protection law, US-American companies who joined the Safe Harbor agreement and are following its principles are still considered to be providing an adequate level of protection. [190] Some national authorities, like the German supervisory authorities, now require a real ("on-the-spot") examination of the validity of the US-company's claim to obey the Safe-Harbor agreement. [191]

[183] *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.
[184] *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.
[185] *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.
[186] C.f. *Art. 29-Working Party*, Working Document Setting up a framework for the structure of Binding Corporate Rules, WP 154.
[187] *Brennscheidt*, Cloud Computing und Datenschutz, p. 173.
[188] *Niemann/Paul*, K&R 2009, 444 (449).
[189] *Brennscheidt*, Cloud Computing und Datenschutz, p. 174
[190] 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
[191] *Düsseldorfer Kreis, Decision from the 28th/29th April 2010, available at:* `http://www.bfdi.bund.de/SharedDocs/Publikationen/Entschliessungssammlung/DuesseldorferKreis/290410_SafeHarbor.pdf?__blob=publicationFile`*; Marnau/Schlehahn, DuD 2011, 311 (315).*

Due to their inability to access their cloud provider's data processor, smaller cloud users may face severe problems to pass these tests. A cloud user's obligation to monitor the cloud provider also contradicts the basic idea of cloud computing - the possibility to outsource the data processing without having to control it anymore. Therefore, compliance by means of the Safe Harbor Agreement seems to be impractical for cloud solutions, at least given the actual practice of some supervisory authorities. [192]

2.4.3.2 The GDPR

Concerning the transfer of data to companies/data processors located outside the EU, the Regulation sticks to the former approach of the DPD. [193] 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 to a third country,thus safeguarding an adequate level of protection (comparable to the European level), which is crucial for any transmission. The instruments which a data processor can use to comply with these requirements remain, basically, the same: The transmission can be based on:

- an acknowledgement of adequacy by the EU Commission (Article 42 GDPR), or
- 'binding corporate rules' (Article 42 (2 a) GDPR), or
- a European Data Protection seal (Art. 42 (2 aa) GDPR) (see 2.3.2.4), or
- standard contract clauses (Art. 42 (2 c) GDPR), or
- contract clauses approved by a supervisory authority (Art. 42 (2 d) GDPR)

Concerning the benchmarks and relevant criteria which the EU Commission will use for acknowledgement of an adequate level, Art. 41 (2) requires the Commission to

> "...give consideration to the following elements:
>
> (a) the rule of law, relevant legislation in force, both general and sectorial, including concerning public security, defense, national security and criminal law as well as the implementation of this legislation, the professional rules and security measures which are complied with in that country or by that international organization, jurisprudential precedents, as well as effective and enforceable rights including effective administrative and judicial redress for data subjects, in particular for those data subjects residing in the Union whose personal data are being transferred;
>
> (b) the existence and effective functioning of one or more independent supervisory authorities in the third country or international organization in question responsible for ensuring compliance with the data protection rules, including sufficient sanctioning powers, for assisting and advising the data subjects in exercising their rights and for co-operation with the supervisory authorities of the Union and of Member States; and
>
> (c) the international commitments the third country or international organization in question has entered into, in particular any legally binding conventions or instruments with respect to the protection of personal data."

[192] *Brennscheidt*, Cloud Computing und Datenschutz, p. 166; *Heidrich/Wegener*, MMR 2010, 803 (806); *Giedke*, Cloud Computing, 233.
[193] *Nebel/Richter*, ZD 2012, 407 (412).

In addition recital 81 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, take into account how a given third country respects the rule of law,"

In a nutshell, the Commission has 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 43 GDPR. BCR have to fulfil certain criteria 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 85 GDPR). BCR will be approved by the Commission if they fulfil Article 43s criteria. The LIBEcommittee of the EU Parliament significaly changed the former version, with regards to cloud computing. BCR shall only be approved by the Commission if they really have a binding character:

> "Article 43 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"

Hence, BCR have to bind not only members of the controller's group but also subcontractors - an amendment which aimed specifically at cloud computing services. [194]
Moreover, the European data protection seal (Article 39) can be used to provide evidence of a processors' compliance with the GDPR by the controller if a processing on his behalf shall take place (see 2.3.2.4). Moreover, the seal can provide evidence for appropriate safeguards concerning the level of data protection in order to permit a transfer to a third country. Hence, the data protection seal can be important for both steps needed to render a third-country transfer lawful.
Appropriate safeguards can also be provided by means of standard contractual clauses or contract clauses approved 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 the European data protection principles. Although the person whose data is being processed is not part of the contract, this person has to be provided with information, according to Article 17 GDPR. Whereas standard contract clauses which are acknowledged by the EU commission benefit of a general validity (Article 62 1 (b) GDPR) individual contract clauses of a controller need a prior authorisation from the competent supervisory authority (not the commission), Article 42 Par 4 GDPR.
The Safe Harbor agreement (see 2.4.3.1) will not be affected by the GDPR. [195] Recital 79 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- ensuring an adequate level of protection for the fundamental rights of citizens"

[194] *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)) 26th of February 2013, p. 140 available at: `http://www.europarl.europa.eu/meetdocs/2009_2014/documents/itre/ad/927/927816/927816en.pdf`.
[195] *Nebel/Richter*, ZD 2012, 407 (412)

Data transfers to a controller or a processor within the USA will, therefore, still be possible if the receiving party follows the Safe Harbor principles. This does not solve the problems arising from the rather self-regulatory character of Safe Harbor described above (2.4.3.1).
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 Art. 44 GDPRs requirements are met. Thus, either the data subject has to give his consent (causing the same problems as described above, see 2.4.2.3) to the transfer or one of the legal permissions in Art. 44 (b) to (g) should be applicable. Those permissions are similar to Art. 6 GDPRs legal permissions (see 2.4.2.2) for processing personal data. Note that Art. 44 takes effect on the second step (if the transfer to a third country is lawful) and not on the first step (if the processing, itself, is lawf

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, Art. 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 [196]. As the DPD does not specify which measures exactly 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. [197]
In order to achieve the goal of enhancing (or guaranteeing) the safety of the 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. [198] 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. [199]
Art 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, as well.
Hence, for PRACTICE, one important tool to ensure confidentiality is encryption.

[196] See also *Art. 29-Working Party*, Opinion 05/2012, WP 196, 14.
[197] 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.
[198] *German Federal Office for Information Security Technology*, Safety Recommendation for Cloud Computing Providers, p. 28 ff.
[199] *Art. 29-Working Party*, Opinion 05/2012, WP 196, p. 14f.

2.4.4.2 Under the GDPR - Article 30 GDPR

The GDPR will change the specification of technical and organisational measures. Article 30 GDPR regulates the controllers and processors duties, regarding the detailed measures to be taken . Nevertheless, the core principles set out in Article 30 are similar to those developed by the DPD. The GDPR includes them explicitly in the norm.

> "Article 30 Par. 1a. GDPR
>
> Having regard to the state of the art and the cost of implementation, such a security policy shall include:
>
> (a) the ability to ensure that the integrity of the personal data is validated;
>
> (b) the ability to ensure the ongoing confidentiality, integrity, availability and resilience of systems and services processing personal data;
>
> (c) the ability to restore the availability and access to data in a timely manner in the event of a physical or technical incident that impacts the availability, integrity and confidentiality of information systems and services"

To determine the state of the art the European Data Protection Board [200] will be entrusted to issue guidelines, recommendations and best practices (Art 30 Par 3 GDPR).
Encryption of data will still be a very useful tool to accomplish the task of ensuring integrity and confidentiality, set by Article 30 GDPR. The measures shall at least,

> "[. . .] protect personal data stored or transmitted against accidental or unlawful destruction, or accidental loss or alteration, and unauthorised or unlawful storage, processing, access or disclosure;" (Art. 30 par. 2 lit. b GDPR)

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" (Art. 30 par. 1 GDPR).

Hence, it will be necessary to monitor the Data Protection Board's publications and always use encryption that is considered as 'State-of-the-art'.

> **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 30 GDPR when using PRACTICE technologies.**

2.5 Other legal changes in the GDPR

2.5.1 Third Country Actions against Data Controllers - Article 43a)

Article 43a Par. 1 provides a verdict on enforceability and "reject-ability" of judgments of a court, a tribunal or a decision of an administrative authority of a third country, regarding the

[200] 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 64 GDPR.

requirement of a controller or processor to disclose personal data. [201] This negative clause clearly aims at activities of third countries obliging providers (data controllers, processors) to disclose personal data - following the NSA scandals and revelations of Edward Snowden. Whereas the first unofficial draft of the regulation by the Commission in late November 2011 (which had been leaked to the public) contained a similar provision in its Article 42, the official proposal in January 2012 omitted that provision. [202]. The EU Parliament reintroduced this Article in obvious reaction to the monitoring activities of foreign intelligence agencies.
Moreover, a US-Court recently obliged a cloud provider to disclose data not only stored in the US but also that which was stored on server based in Ireland. [203] The court denied to quash a warrant after Microsoft filed an instant motion against it. The warrant based on of the US Stored Communications Act obliged Microsoft to disclose information to the US government, and was specifically referring to an e-mail account hosted in Dublin, and, therefore, stored within the EU. Whereas Microsoft argued that the Stored Communications Act cannot be applied extraterritorially, but the court refused to accept this argument and extended the application of the SCA to third countries.
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 90 GDPR states that

> "In cases where controllers or processors are confronted with conflicting compliance requirements between the jurisdiction of the EU, on the one hand, and that of a third country, on the other, the Commission should ensure that EU law takes precedence at all times"

In case of an order issued by a third-country court or supervisory authority, etc,. the controller or processor and, if existing, the controller's representative, shall notify the supervisory authority of the request without undue delay and must obtain prior authorization for the transfer or disclosure by the supervisory authority (Article 43a Par. 2). In principal, no judgment of a court or tribunal of an administrative authority in a third country will be recognized in the EU if a controller or processor is forced to disclose personal data (Article 43a Par. 1). The supervisory authority has to assess the compliance of the requested disclosure with the regulation and, in particular, if the disclosure is necessary and legally required in accordance with Article 44 Par. 1 d and Article 44 Par. 5. Without prejudice to Article 21, the controller or processor must also inform the data subjects of the request and of the authorization by the supervisory authority and, where applicable, inform the data subject whether personal data was provided to public authorities during the last consecutive 12-month period, pursuant to point of Article 14 Par. 1.
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 43 (a). If the supervisory authority does not acknowledge the data transfer required by a third-country authority, according to the GDPR,

[201] *Stadler*, Der Datenschutz bietet keine Handhabe gegen die überwachungspraxis der Geheimdienste; Klinger, jurisPR-ITR 6/2014 annotation 2.
[202] *Logemann*, LIBE-Ausschuss bestätigt Gesetzentwurf zur EU-Datenschutz-Grundverordnung; Bergemann, EU-Datenschutzverordnung darf nicht Merkels NSA-Feigenblatt werden; 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), Version 56 (29/11/2011), available at: `http://statewatch.org/news/2011/dec/eu-com-draft-dp-reg-inter-service-consultation.pdf`
[203] In re Warrant to Search a Certain E-Mail Account Controlled and Maintained by Microsoft Corp., F. Supp. 2d., No. 13 Mag. 2814, 2014 WL 1661004, at *11 (S.D.N.Y. Apr. 25, 2014), CRi 2014, 91 and available at: `http://www.nysd.uscourts.gov/cases/show.php?db=special&id=398.`

the controller or processor will be in a collision of obligations. [204] This difficult situation is addressed in Recital 90 GDPR:

> "The Commission should provide guidance and assistance to the controller and processor, and it should seek to resolve the jurisdictional conflict with the third country in question"

Nevertheless, this declaration of will does not really provide a clear solution for the dilemma of a cloud provider under the control of European law and the law of a third country.

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. [205] 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 the 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. However, the GDPR obliges the controller and the processor to respect the principle of Privacy by Design; in contrast, the developer of a data processing technology is not directly addressed by Art. 23 GDPR.
From a policy perspective, this solution seems to miss the relevant aspects of Privacy by Design, as a controller or a processor is often not be 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 offer their service based on technologies offered by third parties (such as software developers etc.). [206]
Concerning Privacy by Design, encryption technologies, as well as other technologies which guarantee privacy by technological means, are one of the crucial enhancements to which the GDPR refers.
As PRACTICE develops privacy-preserving cloud computing technologies, users are able to use cloud computing for computation over encrypted data. As a benefit from the encryption, the data will be secured against unauthorized access and alteration. The technologies are developed to improve privacy and prevent third parties from learning confidential values from the start. **Thus, the principle of Privacy by Design is fulfilled par excellence.** With those privacy-preserving cloud technologies, controllers and processors can live up to Art. 23 GDPR's requirements.

2.5.3 Right to erasure

Whereas the (heavily-debated) 'right to be forgotten and to erase' affected providers seriously by obliging them to ensure that data would also be deleted on third-party caches and servers, the new proposal of the GDPR (as adopted by the LIBE-committee of the EU Parliament) provides only for a 'right to delete' or erasure in Article 17. [207] Although the term has changed, the

[204] *Plath*, Datenherausgabepflicht für Cloud-Anbieter nach US-Recht vs. EU-Datenschutzrecht.
[205] *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
[206] *Roßnagel/Richter/Nebel*, ZD 2013, 103 (105).
[207] *Fazlioglu*, International Data Privacy Law, 2013, p. 149; *Sartor*, International Data Privacy Law 2013, 3 (9); *Krügel*, ZD-Aktuell 2014, 03870; *Roßnagel/Richter/Nebel*, ZD 2013, 103 (107).

original proposal's content continues to exist. The data controllers should, therefore, be obliged to provide information about a deletion request of an interested party to third parties to whom data has been passed on. However, a lot of details still remain unresolved: for instance, 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 ECJs recent Google Spain decision. Unfortunately, the ECJ stated very briefly that the data subject's rights, as a general rule, override the interest of the public. [208] However, the ECJ also conceded that there are specific cases in which the interest of the person whose data is being processed may be outweighed by (depending on the sensitivity of information stored) the publics' interest inf accessing that information (for instance, in cases of persons of public interest).
For cloud computing it is, therefore, important that the cloud user is able to force 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).

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 EU Commission (and the Parliament) to overcome these deficits is an increase of fines in Article 79 to a maximum of 2% of the global annual turnover of an infringing company - which is comparable only to antitrust fines.
Even more, Article 79 (6) GDPR provides for fines of up to 5% of the annual worldwide turnover of a company, or up to 100 million Euros if severe breaches of data protection duties should occur. [209]
Finally, Article 63 Par 1 GDPR envisages to strengthen enforcement across borders:

> "For the purposes of this Regulation, an enforceable measure of the supervisory authority of one Member State shall be enforced in all Member States concerned."

Thus, the provision introduces some form of mutual acknowledgment of enforceable orders in Europe.

2.5.5 Report of Data Breach, Article 31

In case of data breaches, the controller has to inform the supervisory authority without undue delay. Moreover, the processor of data has to inform the controller without undue delay about any data breaches. [210] 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 31 Par 3 GDPR lists the minimum requirements that a notification has to meet. The notification has to include the 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

[208] ECJ, Judgment of 13th May 2014 in Case C-131/12 - Google Spain SL/Google Inc. v AEPD/Mario Costeja Gonzalez, rec. 81.
[209] *Natz/Wolters*, LMuR 2014, 3 (4); *Krügel*, ZD-Aktuell 2014, 03870.
[210] Regarding the difference between controller and processor see 2.3.2

for cloud computing has been compromised, due to the scalability of cloud computing and the fast transfer of data sets (). [211]
The Data Protection Authority must keep a public register of the types of breaches notified. Article 31 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.

[211] 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

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 out 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]

3.1.1.1.1 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, like 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

[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.

uses symmetric-key encryption, avoids fully homomorphic encryption, and runs on unmodified DBMS software (by using user-defined functions).

3.1.1.1.2 Adjustable Query-Based Encryption

The encodings differ in their encryption and safety. [5] Some ways of encryption are more easily decrypted than others but have to be used for certain queries over the information stored on the DBMS server. To prevent disclosure of all possible encryption of the 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.

3.1.1.1.3 Chain Encryption Keys to User Passwords

One of the principle ideas of Encrypted HANA refers to chaining the encryption. [6] By 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 server depends on the classes of computation required by the application's queries; thus, the DBMS server 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 couldn't 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 Server never receives decryption keys of the plaintext; thus, the

[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.

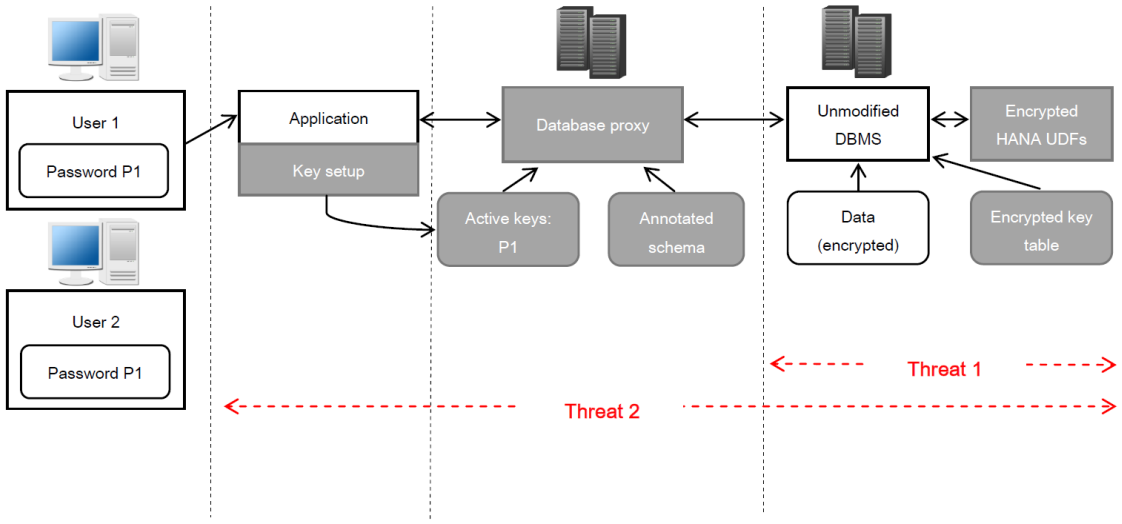

Figure 3.1: Architecture of Encrypted HANA

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 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 returned to the application. Therefore, it is possible that a malicious administrator or external attacker may gain access to the application, the database proxy or the DBMS server, 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 completely 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 server only gets access to an anonymized scheme, encrypted user data, and some auxiliary tables used by Encrypted HANA. Encrypted HANA also supplies the server with Encrypted HANA-specific,

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

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. Afterwards the proxy examine if the DBMS server should be given keys to adjust encryption layers before carry out the query, and if so, issues an UPDATE query at the DBMS server 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 server which carry out it using standard SQL [15]

4. In the end, the DBMS server 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]

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 has to be qualified as "processing" according to the DPD. During this first step, the entity which transfers the data is then qualified as the controller as well as the processor, at the same time. The storage of data is expressly mentioned in Art. 2 (b) of the DPD as an action considered processing.

[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.

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. Running the queries means using, aligning or combining, and consulting the data, according to Art. 2 (b) of the DPD. The controller would still be the client; the processor would now be considered to be the provider of Encrypted HANA, as they would do the actual computation.
Hence, we will have to carefully assess in the following if the DPD is applicable, in particular if personal data is really being affected or sufficiently anonymised 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 server, etc. As shown above, it is controversial which effect encryption has on personal data.
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 server. Hence, regardless of the encryption, personal data is processed by the DBMS server 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. [20] 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 personal data is not affected any more due to encryption. In other words, the applicability of the data protection law depends on the party receiving the data, not the one sending it. [21] 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 less safe 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 - so that the data still remains personal data (for everyone).

[20] 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
[21] *Dammann*, in: Simitis, BDSG, par. 3, recital 34.

3.1.2.3 Compliance with data protection law now and in the future

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 concerning the notion of personal data and the level of encryption. As neither the DPD nor the GDPR implement a clear definition of personal data, and moreover the ECJ still does not give any clue which approach the court does favor, we have to do the analysis in a two-folded way 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 we assume the applicability of the DPD, in spite of encryption, 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 done and had freely given their consent, the controller complies 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, then the processing may take place only in the case of explicit legal permissions such as outweighing interests of the processor or fulfilling contractual obligations, etc. (in the relationship to the data subject - not in relationship to the cloud provider and cloud user). However, this could hardly 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 to 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 where exactly 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 Art. 17 of the DPD ("order processing"). Thus, the controller would be treated as if he were running the Encrypted HANA-technology himself. Therefore, there would be no 'transfer' to another entity when the data is stored on the DBMS-server and, also, no permission would be needed to do so. To run queries over the data, on the other hand, 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.
It should be ensured that reengineering the query-results in order to obtain personal information should be as difficult as possible. If not, providing a query-result would result in transferring personal data to the entity receiving it - thus bringing the DPD into play again.
The Encrypted HANA provider also 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 over encrypted data possible, Encrypted HANA meets

the idea of minimizing the amount of personal data that has to be processed. [22]

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 upcoming regulation does not, unfortunately, solve the dispute concerning the approach on defining 'identifiable' data. However, recital 23 of the proposal for a regulation refers to all means reasonably likely to be used by the controller or by any other person that should be taken into account. Whereas the words 'by any other person' might suggest an absolute approach, it is crucial 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 the used state-of-the-art encryption could be overcome with reasonable efforts. Hence, the storage and computation over 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 Art. 4 par. 2b GDPR. 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). For those who are able to decrypt the data (the Encrypted HANA user), however, is identifiable; therefore, from the perspective of these users, we have to qualify the data as personal data.
This distinction between different parties processing the data is the main difference between the relative and the absolute approaches, the latter do not accept such a distinction. Technically speaking, whereas the DBMS-server provider stores encrypted data, he 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 might gain acceptance. 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 server 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 him should be constituted as an '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 thenhis duties following Art. 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 Art. 28 GDPR demands and oblige the processor to take technical and organizational measures demanded by Art. 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.
A risk analysis of the potential impact of the data processing according to Art 32a GDPR has to be carried out by the controller; moreover, under certain circumstances a data impact

[22] *Schaar*, Privacy by Design

assessment has to be carried out by either the controller or the processor, according to Art 33 GDPR. 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 his duties, the provider the DBMS server should apply for a certification (a privacy seal as described in 2.3.2.4) so that he guarantees 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 his behalf. If the provider has been certified the users of encrypted HANA would not have to monitor and prove the compliance on their own rather than 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 adequacy acknowledgment of the EU Commission (Article 42 GDPR), 'binding corporate rules' (Article 42 (2 a) GDPR), a European Data Protection seal (Art. 42 (2 aa) GDPR) (see 2.3.2.4), standard contract clauses (Art. 42 (2 c) GDPR) or contract clauses approved by a supervisory authority (Art. 42 (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 ground of one of the permissions in Art 44 GDPR. The data protection can also be used in order to provide evidendce for a lawful transfer to a third country.
Hence, the DBMS-server provider's advantage of getting a certification is even greater if he is based in a third county. According to Art. 43a, 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 countries' 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.) turn up. Hence, if a new system using Encrypted HANA is set up, 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 server 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 has to 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.
Before the controller collects his clients' data, he will have to provide the clients with information, according to Art. 14 GDPR. Therefore, he (the controller) will have to inform the clients who his processor (the DBMS-server provider) will be, where the server will be located and for which purposes he intends to process the data, among other information (see 2.4.2.4). If the client exercises his rights according to Art 17 GDPR (the so called 'right to erasure'), the user of Encrypted HANA will have to ensure the complete deletion of the clients data from the DBMS server. If the Art 35 GDPRs 'trigger' of processing data relating to more than 5000 data subjects in a consecutive 12-month period is pulled, the user of Encrypted HANA and the provider of the DBMS-server have to designate a data protection officer (see 2.5.6).

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. [23] 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. [24] 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. [25] Sharemind Server is an application and 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 not collude. The system is capable of performing computations on input data without compromising its privacy. Moreover, it can process data whilst shielding data even from access of the server administrator because the system is based on solid cryptographic foundations.

3.2.1.1.1 Architecture of Sharemind

Sharemind uses Secure Multiparty Computation and secret sharing to protect the personal data of the user. [26]

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 jointly compute a useful result. [27] Every party will provide an input value and learn only the result of his value so that nobody gets the whole information. [28] In the case of data aggregation algorithms, it is generally not possible to learn the inputs of other parties from the result. [29] Figure 1 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 which kind of information the other donor has to provide from its database (for an example the IDs of the persons whose data is about to be compared). The separation of servers between

[23] An overview of the project can be found under: https://sharemind.cyber.ee/introduction-to-sharemind
[24] An overview of the project can be found under: https://sharemind.cyber.ee/introduction-to-sharemind
[25] *Bogdanov*, Sharemind: programmable secure computations with practical applications, p. 30
[26] *Kamm/Willemson*, Secure Floating-Point Arithmetic and Private Satellite Collision Analysis, p. 5.
[27] *Bogdanov*, Sharemind: programmable secure computations with practical applications, p. 24
[28] *Kamm/Willemson* , Secure Floating-Point Arithmetic and Private Satellite Collision Analysis, p. 2.
[29] *Kamm/Willemson* , Secure Floating-Point Arithmetic and Private Satellite Collision Analysis, p. 2.

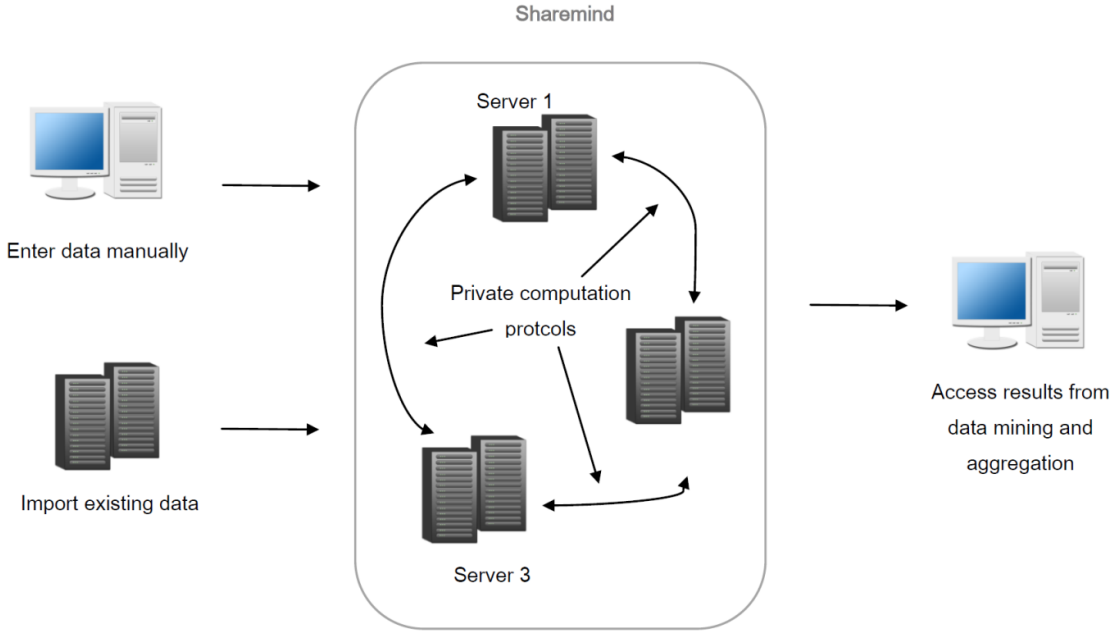

Figure 3.2: Function of Sharemind, three Sharemind Servers were deployed by three independent organizations, the information is divided between the three and every one of them receives a part of the information and works with it. At the end every part sends his results back to the client

input donors and servers is useful as it does not force every party to run secure multiparty computation protocols. [30]
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. [31] 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. [32] 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. [33]

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. [34]

[30] *Kamm/Willemson*, Secure Floating-Point Arithmetic and Private Satellite Collision Analysis, p. 3.
[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. 4.
[34] *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 of prominent use case 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 spacecrafts are flying around the Earth. [35] 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. [36] The two orbital planes intersected at a nearly right angle, resulting in a collision velocity of more than 11 km/s. [37] The locations and orbits of the communication satellites are sensitive information; hence, governments or private enterprises want to protect this data. [38] Whereas 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. [39]
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. [40] 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. [41] Then, the operators secret-share their data and upload the shares to the three servers. Now, collision analysis is a collaborative effort between the three hosting parties and the satellite operator, who can query the results of the analysis. [42]
The same method could be used for many 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 intermediate 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

[35] NASA, NSSDC Master Catalog, last accessed February 16, 2014 available at: `http://nssdc.gsfc.nasa.gov/nmc/`.
[36] NASA, Orbital Debris Quarterly News, 2009, vol. 13, Issue 2, 1 f.
[37] 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.
[38] *Kamm/Willemson*, Secure Floating-Point Arithmetic and Private Satellite Collision Analysis, p. 8.
[39] *Kamm/Willemson*, Secure Floating-Point Arithmetic and Private Satellite Collision Analysis, p. 8
[40] 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.
[41] *Kamm/Willemson*, Secure Floating-Point Arithmetic and Private Satellite Collision Analysis, p. 8.
[42] *Kamm/Willemson*, Secure Floating-Point Arithmetic and Private Satellite Collision Analysis, p. 8.

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 that ID information would need the data subject's consent or an explicit legal permission. Alternatively, all data can be loaded to Sharemind and securely join them using ciphertexts. This would reduce the amount of personal data shared during the work done with Sharemind and, therefore, comply with the principle of Privacy by Design (if the new regulation comes into force). Before the data is stored on the different servers, it has to be divided. This process must be done over 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 Art. 2 (a) of the directive; it is, rather, whether dividing of personal data still has to be seen as processing personal data. In Art. 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. [43] 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 extreme way.

Concering 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 started the research done with Sharemind and decided to use Sharemind. According to the criteria relevant to define the notion 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 done. 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 just started, 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 without discussing 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,

[43] Also *Gola/Schomerus*, Bundesdatenschutzgesetz 2012, par. 3 recital 30.

in the light of the DPD.

3.2.1.2.2 Applicability of data protection law

The advantage of the multiparty computation refers to the use of just random fragments of personal data. The original data is only 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, however, 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 [44], one important parallel may be used: 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. [45]
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 would the directive be applicable. Theoretically, all server providers may collude and reengineer 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 exclude the applicability of the Data Protection Directive. However, 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.
Hence, we have to analyze the legal situation in the sense of a "worst-case scenario" if the absolute approach would prevail (please note that we still believe that the relative approach is the right one)

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

[44] 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.

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

the sense of Art. 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 statistical query, such as students etc., a user can be a research (or commercial) organization that would use the data in order to have them combined with other data which stem from other parties. In such a scenario, the "user" would be qualified as a data controller. Moreover, even another provider may be involved; for instance, a provider of statistical methods and software who determines the data processing; then the issue of a joint controllership is raised again. Even more, 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) has to be bound by contract to refrain from using Sharemind in a way that would give away information about the persons which the data donors provided in the first place (for instance, 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). For the user of Sharemind, 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 - in order to deter the user from re-identification.

Compliance with the GDPR The GDPR will only be applicable to the data processing done 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 not likely that a single part of a secretly-shared date - enables the identification of a person. , as more than a single key is needed to decrypt the date 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 on 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 GDPRs 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 controllers jointly are possible. To ensure a lawful processing under the GDPR all involved processors should process the data on behalf of the respective controller (the user, the client). Hence, before Sharemind is used the involved parties 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 Art. 22 ss. of the GDPR the controller has either to obtain the consent of affected persons or to benefit of an explicit legal permission. Therefore, like under the DPD (see 3.2.1.2.3) it is highly recommended to get the data subjects consent (see 2.4.2.4).
The GDPR addresses, like the DPD, joint controllership. Both controllers are responsible for the use of Sharemind. They will be bound by Article 24 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. 31 GDPR), or eventually (if necessary) the appointment of an data protection officer (2.5.6). 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 Art. 17 GDPR (see 2.5.5).
Sharemind makes it possible for computation to be done over data without the computing parties learning the data information. If it is ensured that the system's output cannot be re-identified without disproportionate efforts, then the goals of Privacy by Design in Art 23 GDPR can be met.

3.2.2 Secure Collaborative Statistics in Credit Rating

3.2.2.1 Functioning

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 business forms). 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, one of the 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 again, 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, in a smaller business scale, too complex to reach.
Another solution is - again - provided by Secure Multiparty Computation 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. MPC does not require one entity to learn all inputs, in contrast to a trusted third party. 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 learning it because only the party holding the data can access it. Only when the secure computation of the benchmark takes place will the data be secretly

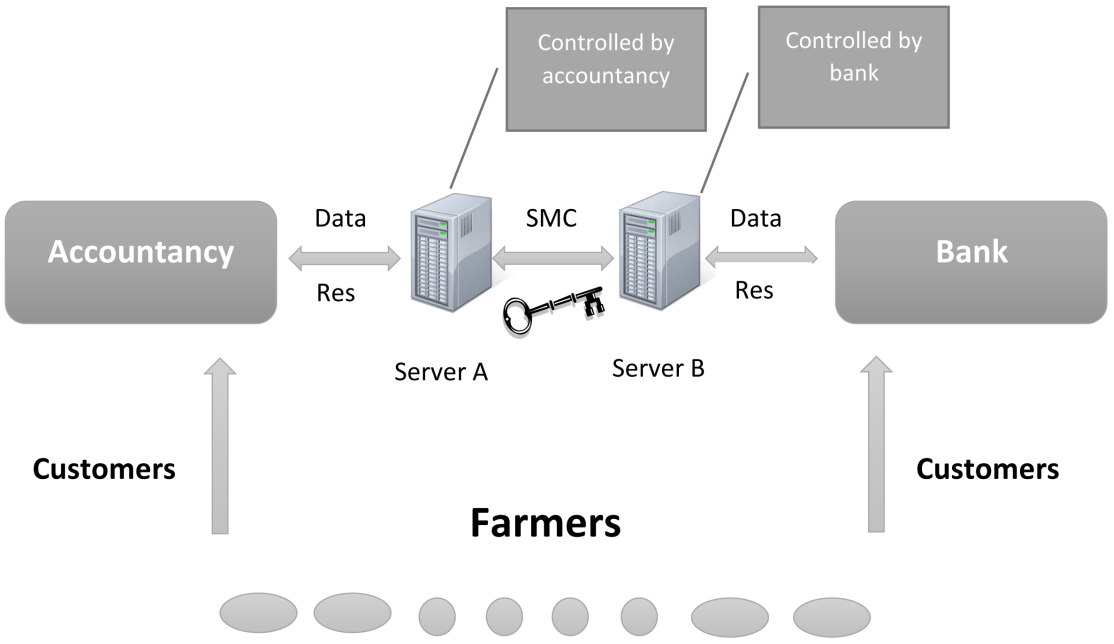

Figure 3.3: The basic functioning of secure collaborative statistics in credit rating using linear programming

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. [46] The bank will be able to get 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 serious lending. 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. [47] 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 indicate the farms this performance analysis refers to.

[46] See `https://cvr.dk/Site/Forms/CMS/DisplayPage.aspx?pageid=21`

[47] 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.

3.2.2.1.3 The possible variations of the system

In the basic 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 other 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 accountancts house and server b run by a bank, the Danish Bankers Association could run server B, so that every bank (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. 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 could 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 date 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.

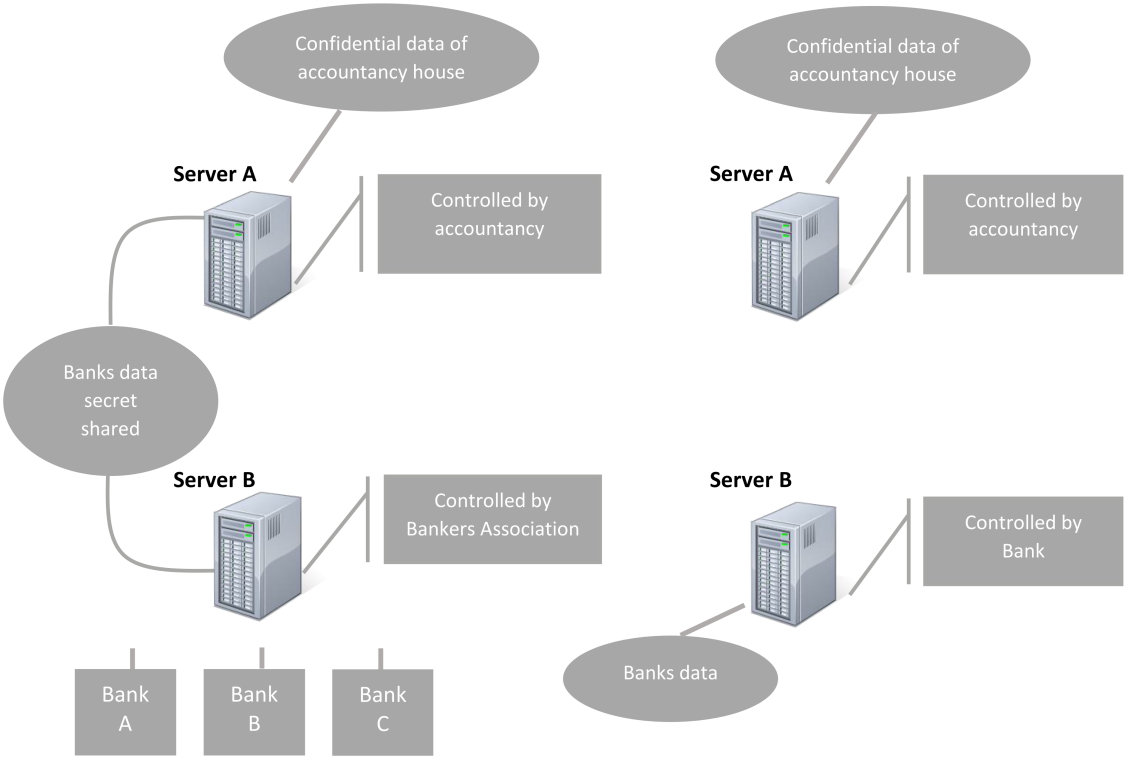

Figure 3.4: Possible Variation of the basic principle by using a third party to host server B

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, the bank, the farms, or, eventually, the Danish Bankers Association and Amazon. The Bank and the accountancy house both determine the purposes and means of the processing, as they decide which data will be computed. Therefore,

they are jointly responsible for the processing of the data as joint controllers, Art 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. Obviously, Amazons does not determine the purposes of the data processing and, therefore, is not qualified as a controller rather than 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 basic 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 house, [48] 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 internally and is not a legal challenge.
If the Danish Bankers Association is involved in a transfer of the bank's data to the server run by the association is needed, then the same criteria and principles Amazon can be applied.
As explained recently the dividing of the data in plaintext cannot be considered 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, too, is a data processing requiring 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, not a company, a corporation or an association. Those are 'legal persons' by the 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 (because there is little to no accountancy data of such farms publicly known the system described might be useful for the financial performance evaluation of such farms). Most of them are privately owned and are not organized as a legal person/entity. However, Denmark opted to include legal persons' data in the scope of its data protection law. [49] Moreover, courts in Germany applied the DPD to legal entities run by a single natural person (one-man-company) when the data reflects information of the natural person. [50] Hence,

[48] Provided that the servers of banks and accountancy firms are based in the EU
[49] *Gola/Schomerus*, Bundesdatenschutzgesetz, par. 3, recital 11.
[50] BGH, NJW 1986, 2505.

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 use 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 parties' 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 accounts the different 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 difference to the Sharemind Case Study refers to the quality of the systems output: whereas with Sharemind, a researcher is not able to re-identify the affected persons and the results are supposed to be anonymous data (see .3.2.1.2.3.1) the use 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 knows which farm is involved these results are clearly personal data. Therefore, the data protection law is applicable on the subject ofcollecting of personal data. If an absolute approach would be taken on 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 data protection law now and in the future

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 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 would need the affected farms' consent or a legal permission for the transfer to Amazon. The consent must be given freely, entangling 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 Art 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 has an interest in only lending money to farms financially stable enough to pay 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 on Amazon.
The same results for the bank (thus reaching the same ends for the bank) could be produced without including Amazon if the bank and the accountancy would run the servers on their own physical machines. For the accountancy, which also would be transferring personal data to a server run on an Amazon the interest would be even weaker compared to the affected farm's

interest, since the accountancy would not obtain a direct advantage (other than making money out of the service or that the bank recognizes the most efficient farms, what would come at the expense of the less efficient farms and therefore might possibly weaken the accountancies interest even more). Thus, it is hard to justify a transfer to Amazon on the grounds of Art 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 been done 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 the 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 safe harbor principles (see 2.4.3.1) - however, entangling all problems already mentioned (2.5.1).
These legal problems can be avoided if the bank and the accountancy 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 useful 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, 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 asthe 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 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). 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. Neither the storage nor the computation of the bank's data should, therefore, need 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 parties' 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 as a processor on behalf of the controller, here the banks. The Association would run one of the servers used as 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 made. 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 an '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 systems 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 has to be considered to be personal data (if the farms' data is considered to be personal, see 3.2.2.2.2). According to the absolute approach (see 2.2.1) the output even 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 done 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 that safeguards have to be taken in order to ensure compliance with the data protection law. Since the output-value is a new data, its production has to be considered to be collection of personal data for the entity retrieving it. For the two entities in charge of producing the output (the accountancy and the bank jointly) the production and the providing 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 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 easier to be obtained from farmers asking for a loan without any cloud-specific problems;unlike in the Sharemind use-case, there is no bigger number of affected individuals whose consent would be needed but one single farm. Even without consent, the collecting of the value might be based on Art 7 lit (b), (f) DPD. Nevertheless, as accountancies do have weaker interests compared to a farms interest in data protection (see above), obtaining consent would be the better 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 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 collecting bank would be illegal, as there is no legal ground for the collection without any contractual performance etc.. The eventual abuse as described 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 use case of Danish banks and farmers implies the storage of data in plaintext and 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) is still be legal on the grounds of the safe-harbor-agreement (see 2.4.2.3). 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. [51] Amazon will do no processing in the described scenario (see 3.2.2.2.3.1)

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

and, therefore, no 'order processing' as regulated in Art. 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. 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 the Art. 22 of the 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 and one bank, or the accountancy and all banks participating, 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). [52]
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 been existent before, but - it will be newly created data. The financial-performance analysis, therefore, will be a collection of data, according to Art. 4 lit (3) GDPR. Hence, the affected farm has to be informed, according to Art. 14 GDPR before the system is put in place. To fulfil Art. 14 GDPRs requirements (especially Art. 14 par. 1 lit (f)), it is once again recommended not to use Amazon EC2 instances to host the two servers.

[52] 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	BetriebsBerater	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
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)
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
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] 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

[3] 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

[4] 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

[5] 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

[6] 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

[7] 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

[8] 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

[9] 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

[10] 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.

[11] Berg, Kay Uwe: EU-Datenschutzgrundverordnung - Das Aus für Auskunfteien und Inkassounternehmen?. In PinG, 69 ff, 2013.

[12] Bergauer, Christian: Indirekt personenbezogene Daten - datenschutzrechtliche Kuriosa. In Jahrbuch Datenschutzrecht, 55 ff, 2011.

[13] 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/

[14] Bitkom: Leitfaden Cloud Computing, 2009. Available at: http://www.bitkom.org/files/documents/BITKOM-Leitfaden-CloudComputing_Web.pdf

[15] 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

[16] Bogdanov, Dan; Kamm, Liiana; Laur, Sven; Pruulmann-Vengerfedt, Pille: Secure multiparty data analysis: end user validation and practical experiments, 2013. Available at: http://eprint.iacr.org/2013/826.pdf

[17] Brennscheid, Kristin: Cloud Computing und Datenschutz, Diss. Bochum, Baden-Baden, 2013.

[18] Charnes, A.; Cooper, W.W; Rhodes, E: Measuring the efficiency of decision making units. In European Journal of Operational Research, 429-444, 1978.

[19] Charnes, A.; Cooper, W.W; Rhodes, E.: Short communication: measuring the efficiency of decision making units. In European Journal of Operational Research, 339, 1978.

[20] 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

[21] Dammann, Ulrich; Simitis, Spiros: EG-Datenschutzrichtlinie - Kommentar. 1st Edition, Baden-Baden, 1997.

[22] 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-deuts

[23] Drews, Stefan; Moneal, Manfred: Grenzenlose Auftragsdatenverarbeitung. In PinG, 143 ff., 2014.

[24] Eckhardt, Jens: Kommentar zu: LG Berlin, Urteil vom 06.09.2007 - 23 S 3/07. In K&R, 601 ff., 2007.

[25] Eckhardt, Jens: IP-Adresse als personenbezogenes Datum - neues Öl ins Feuer. In CR, 339 ff., 2011.

[26] Eckhardt, Jens; Kramer, Rudi; Mester, Brita Alexandra: Auswirkungen der geplanten EU-DS-GVO auf den deutschen Datenschutz. In DUD, 623 ff., 2013.

[27] Ehmann, Eugen; Helfrich, Marcus: EG-Datenschutzrichtlinie - Kurzkommentar. 1st. Edition, Cologne, 1999.

[28] 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

[29] 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

[30] Funke, Michael; Wittmann, Jörn: Cloud Computing - ein klassischer Fall der Auftragsdatenverarbeitung?. In ZD, 221 ff., 2013.

[31] Gerlach, Carsten: Personenbezug von IP-Adressen. In CR, 478 ff., 2013.

[32] German Federal Office for Information Security Technology: BSI-Standard 100-1

[33] 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

[34] 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

[35] Giedke, Anna: Cloud Computing: Eine wirtschaftsrechtliche Analyse mit besonderer Berücksichtigung des Urheberrechts. Diss., Munich, 2013.

[36] Gola, Peter; Schomerus, Rudolf (ed.): BDSG Bundesdatenschutzgesetz - Kommentar. 11th Edition, Munich, 2012.

[37] Härting, Niko: Datenschutzreform in Europa: Einigung im EU-Parlament : Kritische Anmerkungen. In CR, 715 ff., 2013.

[38] Härting, Niko: Internetrecht. 5th Edition, Cologne, 2014.

[39] Härting, Niko: Starke Behörden, schwaches Recht - der neue EU-Datenschutzentwurf. In BB, 459 ff., 2012.

[40] Härting, Niko: Schutz von IP-Adressen. In ITRB, 35 ff., 2009.

[41] 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.

[42] Heidrich, Joerg; Wegener, Christoph: Sichere Datenwolken Cloud Computing und Datenschutz. In MMR, 803, 2010.

[43] 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

[44] Hennrich, Thorsten: Compliance in Clouds. In CR, 546 ff, 2011.

[45] Hilber, Marc: Handbuch Cloud Computing. Cologne, 2014.

[46] 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

[47] 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.

[48] 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

[49] 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

[50] Hornung, Gerrit; Sädtler, Stephan: Europas Wolken - Die Auswirkungen des Entwurfs für eine Datenschutz-Grundverordnung auf das Cloud Computing. In CR, 638 ff., 2012.

[51] Jandt, Silke; Roßnagel, Alexander: Datenschutz in Social Networks - Kollektive Verantwortlichkeit für die Datenverarbeitung. In ZD, 160 ff., 2011.

[52] Jaspers, Andreas: Die EU-Datenschutz-Grundverordnung : Auswirkungen der EU-Datenschutz-Grundverordnung auf die Datenschutzorganisation des Unternehmens. In DuD, 571 ff., 2012.

[53] Jotzo, Florian: Gilt deutsches Datenschutzrecht auch für Google, Facebook & Co. bei grenzüberschreitendem Datenverkehr?. In MMR, 232 ff., 2009.

[54] Kamm, Liina; Willemson, Jan: Secure Floating-Point Arithmetic and Private Satellite Collision Analysis. Available at: http://eprint.iacr.org/2013/850.pdf

[55] Kilian, Wolfgang; Heussen, Benno (ed.): Computerrechts-Handbuch: Computertechnologie in der Rechts- und Wirtschaftspraxis. Supplement 32, Munich, 2013.

[56] Klar, Manuel: Räumliche Anwendbarkeit des (europäischen) Datenschutzrechts. In ZD, 109 ff., 2013.

[57] 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.

[58] 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

[59] Kos, Clemens; Englisch, Bastian: Auftragsdatenverarbeitung? - Gegenüberstellung der aktuellen Rechtslage und der DS-GVO in der Fassung des LIBE-Entwurfs. In ZD, 276 ff., 2014.

[60] 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

[61] Kroschwald, Steffen: Verschlüsseltes Cloud Computing : Auswirkung der Kryptografie auf den Personenbezug in der Cloud. In ZD, 75 ff., 2014.

[62] Krügel, Tina: LIBE-Kompromissvorschlag zur DS-GVO. In ZD-Aktuell, 03870, 2014.

[63] Kühling, Jürgen: Auf dem Weg zum vollharmonisierten Datenschutz?!. In EuZW, 281 ff., 2012.

[64] Kühling, Jürgen; Klar, Manuel: Unsicherheitsfaktor Datenschutzrecht - Das Beispiel des Personenbezugs und der Anonymität. In NJW, 3611 ff., 2013.

[65] Kuner, Christopher: European Data Protection Law. 2nd Edition, New York, 2007.

[66] 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

[67] Leutheusser-Schnarrenberger, Sabine: Zur Reform des europäischen Datenschutzrechts. In MMR, 709 f., 2012.

[68] Maisch, Michael Marc: Nutzertracking im Internet. In ITRB, 13 ff., 2011.

[69] 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

[70] Marnau, Ninja; Schlehahn, Eva: Cloud Computing und Safe Harbor. In DuD, 311 ff.; 2011.

[71] Meyerdierks, Per: Sind IP-Adressen personenbezogene Daten?. In MMR, 8 ff., 2009.

[72] Millard, Christopher: Cloud Computing, Oxford, 2013.

[73] Nägele, Thomas; Jacobs, Sven: Rechtsfragen des Cloud Computing. In ZUM, 281 ff., 2010.

[74] 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

[75] Natz, Alexander; Wolters, Chris: Geplante EU-Datenschutz-Verordnung: Auswirkungen für die Datenverarbeitung in Unternehmen. In LMuR, 3 ff., 2014.

[76] Nebel, Maxi, Richter, Philipp. Datenschutz bei Internetdiensten nach der DS-GVO - Vergleich der deutschen Rechtslage mit dem Kommissionsentwurf. In ZD, 407 ff., 2012.

[77] Niemann, Fabian; Paul, Jörg-Alexander: Bewölkt oder wolkenlos - rechtliche Herausforderungen des Cloud Computings. In K&R, 444 ff., 2009.

[78] Niemann, Fabian; Ammann, Jörg-Alexander: Praxishandbuch Rechtsfragen des Cloud Computing. Berlin, Bosten, 2014.Nord, Jantina; Manzel, Manzel: Datenschutzerklärungen- - misslungene Erlaubnisklauseln zur Datennutzung : -Happy-Digits- und die bedenklichen Folgen im E-Commerce. In NJW, 3756, 2010.

[79] Pahlen-Brandt, Ingrid: Datenschutz braucht scharfe Instrumente: Beitrag zur Diskussion um -personenbezogene Daten. In DuD, 34 ff., 2008.

[80] Peifer, Karl-Nikolaus: Verhaltensorientierte Nutzeransprache - Tod durch Datenschutz oder Moderation durch das Recht?. In K&R, 543 ff., 2011.

[81] 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/

[82] 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/

[83] Pohle, Jan; Ammann, Thorsten: Software as a Service - auch rechtlich eine Evolution?. In K&R, 625 ff., 2009.

[84] Pollirer, Hans-Jürgen; Weiss, Ernst M.; Knyrim, Rainer: Datenschutzgesetz 2000 (DSG 2000) samt ausführlichen Erläuterungen. 2nd Edition, Vienna, 2014.

[85] Popa, Raluca Ada: Research Statement. Available at: http://www.mit.edu/~ralucap/researchstatement.pdf

[86] 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

[87] 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

[88] Rammos, Thanos: Datenschutzrechtliche Aspekte verschiedener Arten "verhaltensbezogener" Onlinewerbung. In K&R, 692 ff., 2011.

[89] Rath, Michael; Rothe, Britta: Cloud Computing: Ein datenschutzrechtliches Update. In K&R, 623 ff., 2013.

[90] Roßnagel, Alexander (ed.): Beck'scher Kommentar zum Recht der Telemediendienste. Munich, 2013.

[91] Roßnagel, Alexander (ed.): Handbuch Datenschutzrecht: Die neuen Grundlagen für Wirtschaft und Verwaltung. Munich, 2003.

[92] Roßnagel, Alexander; Scholz, Philip: Datenschutz durch Anonymität und Pseudonymität : Rechtsfolgen der Verwendung anonymer und pseudonymer Daten. In MMR, 721 ff., 2000.

[93] Roßnagel, Alexander; Richter,Philipp; Nebel, Maxi: Besserer Internetdatenschutz für Europa - Vorschläge zur Spezifizierung der DS-GVO. In ZD, 103 ff., 2013.

[94] Sartor, Giovanni: Providers' liabilities in the new EU Data Protection : Regulation: A threat to Internet freedoms?. In International Data Privacy Law, 3 ff., 2013. Available at: http://idpl.oxfordjournals.org/content/3/1/3.full.pdf+html

[95] Schaar, Peter: Privacy By Design. Available at: http://www.bfdi.bund.de/SharedDocs/Publikationen/EN/0610EUPrivacyByDesign.pdf?__blob=publicationFile Schneider, Jochen: Handbuch des EDV-Rechts. 4th Edition, Cologne, 2009.

[96] Simitis, Spiros (ed.): Bundesdatenschutzgesetz Kommentar. 7th Edition, Baden-Baden, 2011.

[97] Spies, Axel: Cloud Computing: Keine personenbezogenen Daten bei Verschlüsselung. In MMR-Aktuell, 313727, 2011.

[98] Spindler, Gerald: Persönlichkeitsschutz im Internet - Anforderungen und Grenzen einer Regulierung. In Verhandlungen des 69. Deutschen Juristentages, Band I Gutachten, Munich, 2012.

[99] Spindler, Gerald: Persönlichkeitsrecht und Datenschutz im Internet - Anforderungen und Grenzen einer Regulierung. In NJW-Beilage, 98 ff., 2012.

[100] Spindler, Gerald: Datenschutz- und Persönlichkeitsrechte im Internet - der Rahmen für Forschungsaufgaben und Reformbedarf. In GRUR, 996 ff., 2013.

[101] Spindler, Gerald: Datenschutz- und Persönlichkeitsrechte im Internet - Der Rahmen für Forschungsaufgaben und Reformbedarf. In GRUR-Beilage, 101 ff., 2014.

[102] 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.

[103] Spindler, Gerald; Schuster, Fabian (eds.): Recht der elektronischen Medien. 2nd Edition, Munich 2011.

[104] 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

[105] Stadler, Thomas: Der Datenschutz bietet keine Handhabe gegen die überwachungspraxis der Geheimdienste. 05/11/2013. Available at: http://www.internet-law.de/2013/11/der-datenschutz-bietet-keine-handhabe-gegen-die-ueberwachungspraxis-der-geheimdiens
html

[106] Sydow, Gernot; Kring, Markus: Die Datenschutzgrundverordnung zwischen Technikneutralität und Technikbezug - Konkurrierende Leitbilder für den europäischen Rechtsrahmen. In ZD, 271 ff., 2014.

[107] Taeger, Jürgen; Gabel, Detlev (ed.): Kommentar zum BDSG und zu den Datenschutzvorschriften des TKG und TMG. 2nd Edition, Frankfurt/Main, 2013.

[108] 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

[109] Voigt, Paul: Datenschutz bei Google. In MMR, 377 ff., 2009.

[110] Weichert, Thilo: Cloud Computing und Datenschutz. In DuD, 679 ff., 2010.

[111] Wieczorek, Mirko: Der räumliche Anwendungsbereich der EU-Datenschutz-Grundverordnung. In DuD, 644 ff., 2013.

[112] Wolff, Amadeus; Brink, Stefan: Datenschutz in Bund und Ländern - Kommentar. Munich; 2013.

Part II

Part II - Risk Assessment

Chapter 4

Analysis and management of risk for cloud-based processes

4.1 Background and motivation

As always it happens when new technologies are introduced, organizations shifting towards the cloud computing paradigm should carefully consider all the associated risks, defining acceptable scenarios and necessary controls before moving part or all of their business. Cloud-computing frameworks have IT risks in common with any externally provided service, but there are also some essential characteristics that require different risk evaluation methods.
From the economic standpoint, the *risk* of an event E for a given actor A is often represented as the product of the probability that E might happen, times the damage (expressed in currency units) to A if E really happened. In symbols:

$$R(A, E) = P(E) * I_A(E) \tag{4.1}$$

In a computer security context, one needs to handle the problem of identifying E (potential security threats) and estimating $I_A(E)$ and $P(E)$. As far as threat analysis is concerned, it is customary to focus on one or more specific business processes. The risk analyst puts herself in the place of a specific actor (e.g. the *process owner*, i.e. the actor in whose interest the process is executed) and asks the following questions [68]:

- *Threat Categorization*: What can happen to information assets involved in a process?
- *Threat Impact*: How severe could that be?
- *(Frequentistic) Threat Probability*: How often might that happen?

Once threats have been identified and Eq. 4.1 has been computed for all of them, the definition of a mitigation strategy can follow. Typically, it involves answering the following questions:

- *Mitigation Effectiveness*: What can be done to reduce the risk?
- *Mitigation Cost*: What does risk mitigation incur?
- *Mitigation Cost/Benefit Ratio*: Is mitigation cost effective?

Accurately quantifying impact is often a challenge, as losses deriving, say, from decreased consumer trust after a security breach can only be estimated in the long run. As far as probability is concerned, there are cases where a frequency-based probability can be assigned to the events, and other cases where this is too difficult or misleading. In the latter case, we will use a notion of *perceived probability* or *belief* based on specific assumptions about actor A's mindset.
Whatever the impact and probability assessment methodology, however, risk analysis has traditionally focused on composing via a suitable aggregator (in the simplest case, a summation) risks $R(A, E_i)$, $i = 1, \ldots, n$ for all known negative events that may affect a specific actor A, considering neither the potential gains that those events may bring to other stakeholders, nor the influence of those potential gains on the probability of E_i, when the event is an attack (or, like disclosure, a certain consequence thereof) and attackers may coincide with such stakeholders.
Business processes often involve storing or transmitting data that is subject to regulatory and compliance requirements. When data falls under regulatory or compliance restrictions, the choice of deploying the process on a cloud hinges on being convinced that the cloud provider is fully compliant with regulations. Otherwise, the process owner risks violating privacy, regulatory or other legal requirements. Any obligation regarding secure data management usually falls on the tenant or user. The implications for maintaining information security are significant when it comes to privacy, business and national security. For cloud tenants with legal privacy obligations, private information must be handled in the same way regardless of using cloud or traditional storage. If a highly regulated process (e.g. a e-health or e-government one) is to take place on a public cloud, then that deployment must fully meet the interests of the tenant and all applicable regulations and laws regarding data confidentiality and leakage prevention.
In this deliverable we focus on a specific category of data leakage events, the ones that bring some party taking part in a collaborative business process to know more information than the process would entail. This may happen due to intentional publishing of supposedly protected information items by other parties, or to carelessness in the protocol implementation and deployment, e.g. when one party is using the same mobile terminal previously used by another and can reconstruct the information items held.
We call these events *process-related data leakages*, in order to distinguish them from other types of eavesdropping attacks. We propose a simple methodology to analyze risk of process-related data leakages; then, we extend our risk analysis methodology to multi-party SMC processes taking place on clouds. We show that our extension is compatible with all methodological choices for the assessment of probability and impact. We will show how mixing outsourcing in a public cloud for non-sensitive data and reserving trusted systems for sensitive data can reduce risk associated to data leakage threats in cloud-based business processes. Classic quantitative approaches proposed in the Seventies are based on estimating threat probabilities as frequencies using statistical information [45]. In our methodology, we rely on the knowledge of the business process model and its underlying micro-economics to attach probabilities to actors misbehavior/violation of confidentiality and to provide an evaluation of costs taking into account the value of disclosed information. Our methodology takes into account the importance of our information assets together with the cloud's deployment and service models.

4.1.1 State of the art on Risk Management

4.1.1.1 Risk management

In general guidelines for *Risk Management* have been established in ISO 31000:2009 [31], and are related to the process whereby organizations treat, in a methodical way, generic risks related

with their activities not specific to any industry or sector. According to the ISO document, a risk management framework should assist an organization in managing its risks and ensure that risk information derived from these processes is adequately used as a basis for decision making and accountability. The risk management process, depicted in Fig 4.1, includes five activities: communication and consultation, establishing the context, risk assessment, risk treatment, monitoring and review. Communication and consultation with internal and external stakeholders are necessary activities to address issues related to the risk itself, and ensure that consequent decisions are taken. By establishing the context the organization selects the internal and external parameters considered when managing risk, and setting the scope and risk criteria for the remaining process. The organizations monitoring and review processes should encompass all aspects of the risk management process, be internally or externally reported, and should also be used as an input to the review of the risk management framework. Risk assessment is the overall process of risk identification, risk analysis and risk evaluation, where the identification refers to the recognition of sources of risk, areas of impacts, events and their causes and their potential; risk analysis involves the determination of the consequences and their likelihood, and other attributes of the risk; risk evaluation is based on the outcomes of risk analysis, and aims to prioritize treatment implementation. In formal risk framework, risk is formally defined as a set of ordered pairs of outcomes (O) and their associated likelihoods (L) of occurrence

$$Risk \equiv \{(L_1, O_1), \ldots (L_i, O_I), \ldots, (L_n, O_N)\}$$

Risk assessment process has been the focus of another ISO document [32], where most diffused techniques are listed and analyzed, and guidance for their selection and application provided. Modern risk assessment has been used in the nuclear power industry, where risk assessment methodologies have been developed to carefully analyze the operations of the potentially dangerous nuclear power facilities. Such methodologies adopted a fault/event trees representation in order to capture all possible plant failure modes graphically.

In 1979, the National Bureau of Standards (then absorbed into the NIST) published its Federal Information Processing Standard (FIPS) 65, Guideline for Automatic Data Processing Risk Analysis [45], introducing the risk assessment standard for large data-processing centers and proposing a new metric for computer-related risks: Annual Loss Expectancy (ALE) $ALE = \sum_{i=1}^{n} I(O_i) \cdot F_i$ where $\{O_1, \ldots, O_N\}$ is the Set of Harmful Outcomes, $I(O_i)$ represents the Impact of Outcome i in dollars, and F_i is the Frequency of Outcome i. In the subsequent years,

Risk assessment methodologies have become a standard practice aimed to let organizations determine and demonstrate their privacy, security, and compliance with other policies to avoid any loss.

Risk assessments all require similar steps to be carried out, including system characterization, threat assessment, vulnerability analysis, impact analysis, and risk determination [15, 34]. Figure 4.2 shows the steps used in the National Institute of Standards in Technology (NIST) risk assessment methodology [46].

In the literature, assessment-based approaches to risk evaluation have been the source of debate ranging from theories about the value of quantitative versus qualitative methods and to internal versus external analysis of the target system. More in detail there are basically three primary methods for risk assessment methods. *Qualitative* methods are based on simple evaluations, not needing to determine numerically the value of the company's assets at risk and threat frequencies. Qualitative methods enable the classification of risks and show the most important area for improvement, missing however a practical evaluation of the probabilities and impacts of risks. On the opposite, *quantitative* methods are based on the numerical values assigned to impact and likelihood of risks. The advantage of quantitative methods is that at the end

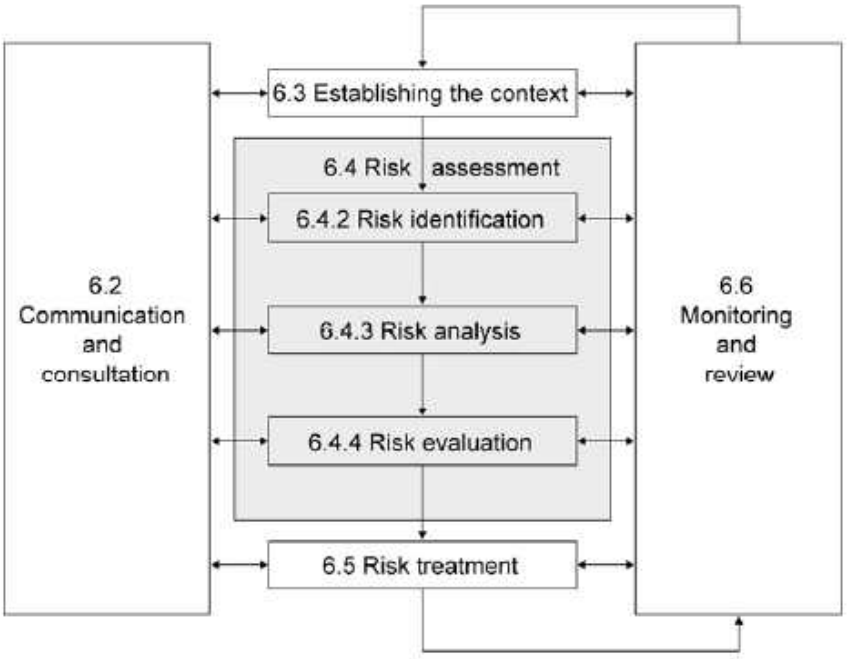

Figure 4.1: Risk Management process in ISO 31000

they provide accurate measurements of risk and impacts, even if a qualitative interpretation of the provided numbers is required. In the middle, *semi-quantitative* (or hybrid) methods are less numerically intensive than quantitative methods and risks are prioritized according to consequences and foreseen probabilities. For these reasons semi-quantitative methods basically take profit of both aforesaid advantages and, therefore, provide risk prioritizations and useful quantifiable impacts analysis.

4.1.1.2 Risk assessment on cloud computing

Cloud computing is a style of computing where "massively scalable IT-enabled capabilities are delivered 'as a service' to external customers using Internet technologies" [26]. If on the one hand the adoption of such a model provides cost savings through economies of scale, on the other hand it introduces some peculiar risk challenges that increase the risks traditionally introduced by any externally provided IT service. Indeed, from the perspective of the security analyst, the delivered services are outsourced in the least transparent way, since data are stored and processed in unspecified servers located in some places, out of the control of the data owner, and shared among different untrusted customers. For these reasons risk management and assessment methods are needed and some researchers have started introducing tailored procedures and techniques to deal with the specific cloud related issues [22, 53].
Various bodies such as the Cloud Security Alliance (CSA), the European Network and Information Security Agency (ENISA), and the US National Institute of Standards and Technology (NIST) have released several documents assisting organizations and customers in the evaluation of the security issues related to cloud computing [17, 13, 46]. The Cloud Controls Matrix released by CSA provides an useful description of the security principles aiming to guide cloud vendors and help cloud clients in assessing overall security risks of a cloud service provider [17]. NIST Special Publication 800-144 provides an overview of the security and privacy challenges

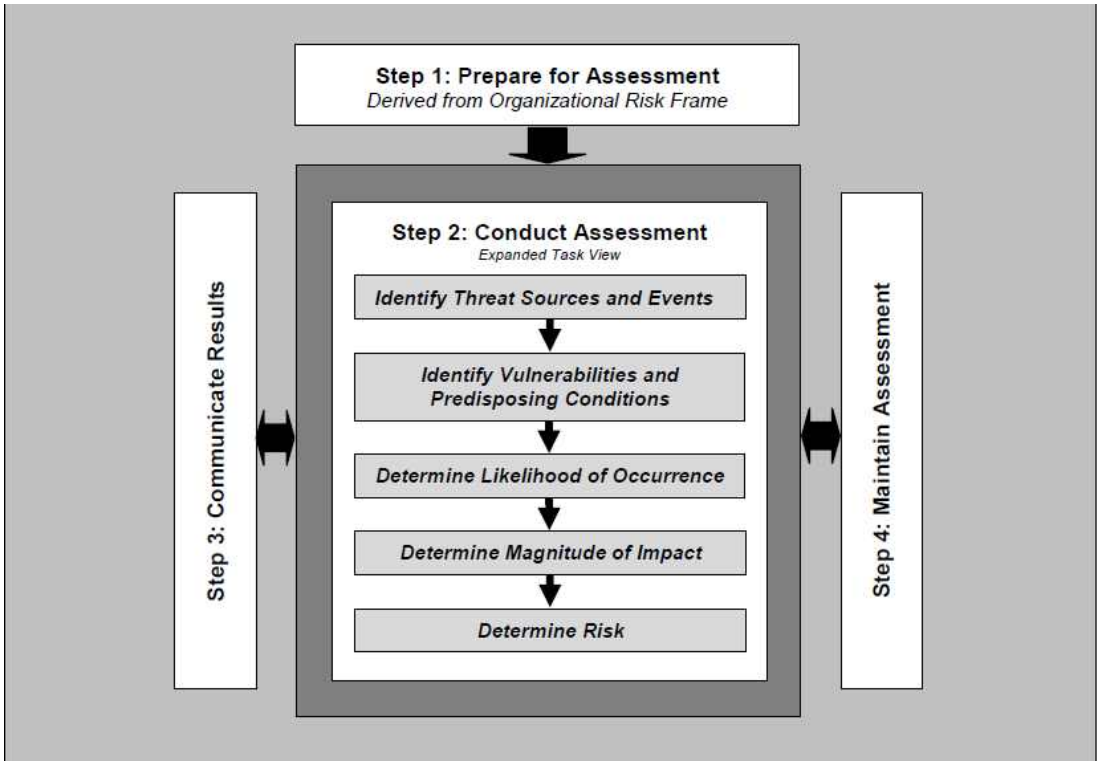

Figure 4.2: Risk assessment process

for public cloud computing and gives recommendations that organizations should consider when outsourcing data, applications, and infrastructure to a public cloud environment.
In [13], the risk assessment process is described using three use-cases scenarios: An SME perspective on Cloud Computing, the Impact of Cloud Computing on service resilience and a scenario involving eHealth. Taking the perspective of estimating the level of risk mapped against the impact, the document proposes a table showing the distribution of the risk probabilities and impacts adopting a scale of 0 to 8, and classifying as low risk values from 0 to 2, medium risk from 3 to 5 and high from 6 To 8. The risks identified in the assessment are classified into three categories: policy and organizational, technical, and legal. Among the first category, lock in, loss of governance, and compliance challenges are classified as high risk having in some cases very high or medium impact. *Lock in* refers to the way tools, procedures and standard data formats or services interfaces are provided on the cloud, verifying that data and service portability are not guaranteed, and the migration for one provider to another may be extremely difficulty for a customer. For this reason a strong dependency for service provision is created on a particular cloud provider, and this could be the cause for a business failure should the cloud provider go bankrupt or be acquired by another company leading to sudden changes in provider policy and non-binding agreements. The loss of governance and control over the data and services could have a potentially severe impact on the organizations mission, potentially leading to the impossibility of satisfying the security requirements, about the confidentiality, integrity and availability of data, and the loss on the control of the performance and quality of offered services. The migration of services to the cloud could also lead to problems related to compliance, since it is difficult for the cloud providers to provide evidence of their own compliance to the relevant requirements, meeting industry standards or regulatory requirements. Other risks related to policy and organizational issues are the possible loss of business reputation due to co-tenant activities, since in cloud computing resources are shared with other tenants, whose

malicious activities may affect the reputation of the other customers who are using the same cloud infrastructure.
Among the technical risks, isolation failure, and malicious insider including the cloud provider itself are classified as high. Isolation failure refers to the failure of mechanisms separating storage, memory and routing between different tenants of the shared infrastructure that can be caused by different kind of attacks, such as the so-called guest-hopping attacks, or SQL injection attacks exposing multiple customers data stored in the same table, and side channel attacks. The malicious activities of an insider could have a severe impact on the security of company's data and services, since in cloud architectures certain roles are needed, such as the cloud provider system administrators and auditors, who can overcome the security barriers or become targets for criminal gangs. Other technical risks such as the compromising of the management interface, or data leakage or distributed denial of service, and other attacks, related to the nature of on-demand distributed and remotely accessed cloud services, have received a medium evaluation.
Legal risks have received almost all an high evaluation. The risk from changes of jurisdiction affects the disclosure of customers' data, since they can be held in centers located in high-risk countries, where the behavior of the legal framework and enforcement can be unpredictable, and the respect of international agreements not guaranteed. In general legal risks coming from breaking compliance with data protection regulation, are rated high since the cloud customer, in this case playing the role of data controller, cannot be sure that the data is managed in a lawful way. Stored customers' data maybe be at high risk of disclosure to unwanted parties, if a confiscation of physical hardware event occurs (as a result of subpoena by law-enforcement agencies or civil suits).
In Atos's white paper [2], the description of a qualitative cloud-specific risk assessment methodology starts by considering the risk factors that are changed when an organization shifts from a traditional infrastructure to a cloud-computing based one. The analysis is based on the risk taxonomy presented by the Open Group [25] and here reported in Figure 4.3.

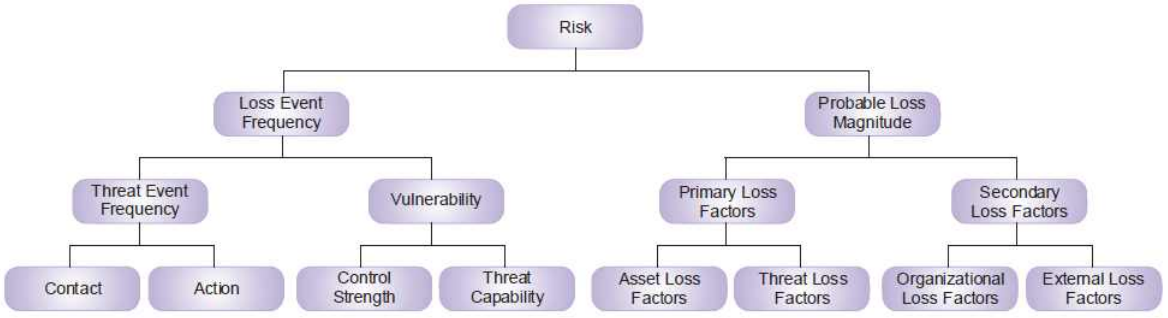

Figure 4.3: The risk taxonomy presented by the Open Group

The transition to a cloud infrastructure may change the probability of the occurrence of a harmful event, reducing the effort necessary to carry on an attack when a cloud specific vulnerability can be exploited. To denote a vulnerability as cloud specific, four indicators have been found. A first category collects all the vulnerabilities that are intrinsic to a core technology of Cloud Computing, such as the possibility for an attacker to escape from the virtualized environment, the possibility to ride or hijack sessions in shared web applications, or the treats to the integrity and confidentiality of data caused by the insecure usage of cryptography or the selection of flawed implementation of cryptographic primitives. Other vulnerabilities have their roots in some essential characteristics of cloud computing, such as ubiquitous network

access, resource pooling, elasticity, on demand self-service, and measured services. Access to the management interface from unauthorized users, attack to the intranet protocols, access to the data not correctly deleted after a resource reallocation, manipulation of metering or billing data, are all examples of vulnerabilities relevant for cloud systems. Other specific vulnerabilities are the ones concerning problems with standard security controls, such as the difficulties to perform network controls in virtualized environment, poor management or storage of the many different kinds of keys needed in a could infrastructure, the difficulty of establishing and using security metrics that are not adapted to monitor the security status of cloud resources. Finally, last category of vulnerabilities are related to cloud offerings, such as the usage of weak authentication schemes, or the exposition to injection attacks (SQL, cross scripting).
The qualitative risk assessment methodology presented in the paper relies on the identification of the security controls that can be used to counter these vulnerabilities, on the basis of known mappings such as the one reported in the NIST report SP800-53 [46] on recommended security controls for federal information systems or the information security standard ISO 27002 [30].
A case study about risk assessment in a cloud computing scenario is offered in [24], where the case of a software company developing business software and adopting a IAAS provided by another cloud service provider is analyzed. The methodology is based on the Risk IT framework, which provides a detailed process model for the management of IT-related risk and on the COBIT 5 framework that assists enterprises in achieving their objectives for the governance and management of enterprise information technology (IT), both provided by Information Systems Audit and Control Association (ISACA) [3]. In the RISK IT framework a list of generic high-level risk scenarios is reported, and a mapping between those scenarios and more general COBIT control objectives is provided, so that a map of risks showing the impact/magnitude and likelihood/frequency of key risks can be created. Based on the map and on the prioritization of the risks, a risk mitigation approach can be adopted, balancing the benefit from ensuring controls and the costs necessary for their implementation.
Some initial work toward a quantitative risk and impact assessment framework for cloud computing, called QUIRC, has been presented in [54]. The QUIRC framework classically defines risk as a combination of (a rough estimate of) the probability of a feared event and its severity, measured as its impact. QUIRC lists six key *Security Objectives* (SO) for cloud platforms, claiming that most of the typical attack vectors and feared events map to one of these six categories. QUIRC's strong point is its fully quantitative approach, which enables stakeholders to comparatively assess the robustness of cloud vendor offerings. However, lack of reliable data on the occurrences of cloud threats in many vertical domains can make QUIRC probability assignment (and the entire notion of "typical" attack vectors) somewhat arbitrary. Another quantitative framework for assessing some security risks associated with cloud computing platforms has been proposed in [58]. The model relies on a fuzzy decision making technique, that allows the definition of the weights of the coefficients for the basic security properties (CIA - Confidentiality, Integrity, and Availability) and the corresponding values of assets relevant for the project, using the knowledge of experts. Then, vulnerability indices are defined for each asset separately and a final fuzzy model is created to compute the impact of each identified risk as product of asset values, vulnerability and threat effects. Even if the resulting prioritization of the risks is valuable, this approach only considers threats to CIA properties. Also, it relies on subjective assessments of likelihood and severity by experts that may be difficult to replicate in practice. Focusing on the same small set of security properties, Khan et al [35] introduced a more systematic approach combining existing tools and techniques such as CORAS [19], and the IRAM (Information Risk Analysis Methodology) with the Threat and Vulnerability Assessment tool (T&VA) [1]. Their technique uses a list of threats provided by the Information

Security Forum (ISF). Depending on the priority of the assets and on the perceived likelihood of the ISF threats, they construct an evaluation matrix and use it to rate the threats' impact on the business. Due to the anecdotal nature of the ISF threat list, whose entries often highlight new and emerging threats rather than frequent ones, this technique can be considered a semi-quantitative one.

4.2 Integration with Privacy Risk Management frameworks

Business processes involving personal data present a inherently high liability due to the risks brought upon the process owner (*controller* in legal terminology) and, possibly, on other stakeholders by violations of the privacy of third parties (*data subjects*). As discussed at length in the first part of this deliverable, a special regulatory framework for personal data processing is currently in force at the European level, prescribing - among other things - that the purposes of the business process involving personal data are clearly defined, that personal data are relevant to such purposes, that personal data are erased at the end of a given time, and that all data subjects have the opportunity to exercise their rights (such as opposition, access, rectification and deletion of their personal data). In addition, the controller has an obligation to take all useful precautions in order to ensure the security of the personal data he processes.
Privacy authorities and regulators have been devoting a huge effort to developing *Privacy Risk Management* (PRM) frameworks [14].
As observed in [38], currently there is still a lack of a systematic approach aimed to provide organization with privacy-supportive processes and practices in the products and services they develop. Towards these objectives, some efforts are conducted in several research projects and common initiatives (such as in the FP7 PRIPARE (PReparing Industry to Privacy-by-design by supporting its Application in REsearch) and PIAF (A Privacy Impact Assessment Framework for data protection and privacy rights)) where the development of new privacy enhancing methodologies and frameworks are addressed.
A privacy impact assessment (PIA) is one of the possible approaches to manage the information life cycle and reduce the privacy risks [69], and for these reasons several government agencies in the UK, Canada and the US require a mandatory PIA. The need for privacy protection has been also recognized by the new European regulation, which promotes the inclusion of Privacy Impact Assessments and Privacy by Design practices in business processes [38], ensuring that all the stakeholders adopt *Privacy by Design* and *Security by Design* methodologies. Introducing these practice should ensure that systems are engineered taking into account privacy and countermeasures to protect IT resources along the development process.
According to recent studies [69], threat analysis should begin at the earliest possible stage of the lifecycle of any business process involving personal data, when there are more opportunities to influence the business process' implementation; it should continue along the business process lifecycle[1]. Here we highlight the potential synergy between our process-oriented, quantitative risk assessment techniques, to be presented in the next Sections, and PRM frameworks.

In principle, privacy risks may be subject to all the forms of analysis introduced in previous Sections. *Qualitative Analysis* uses ordinal scales expressed in words to quickly assess the

[1]PRM is also an important part of a PIA (Privacy Impact Assessment), *a systematic process for evaluating the potential effects on privacy of a process, initiative or proposed system or scheme and finding ways to mitigate or avoid any adverse effects* (http://www.piafproject.org).

relative severity of risks. This technique is often used when numerical data is not available and/or the process targeted by privacy risk assessment is only partially known by the risk assessor, as is often the case at early design stage[2]. *Semi-Quantitative Analysis* adds a score expressed in points to the ordinal scale (e.g. "1" = Low; "5" = Very High)[3]. In privacy risk analysis, a common understanding is that values should not be used to compare relative risks (i.e. the impact of one risk vs. another). Finally, *Quantitative Analysis* computes comparable numeric values to assess impacts (expressed in monetary terms) and probabilities of privacy threats.
The quantitative approach is generally more complex to undertake, requiring full knowledge of the business processes to be analyzed and, in many cases, the development of organization-specific value models to assess the value of disclosed information as seen by different actors. For this reason, the quantitative approach is only applied to evaluate the most significant risks. In the next Sections, we will focus on the design of a method for quantitative assessment of specific categories of privacy risks. Three basic criteria underly our design:

1. Process-orientation: our technique assesses impacts and probabilities of disclosure threats for each specific business process, based on full knowledge of the process model and of the micro-economics underlying it (expressed in monetary terms). Such knowledge is extracted for the target process families identified by the PRACTICE project. As we shall see, process orientation allows privacy risk to be assessed by a number of organizational roles that qualify as process owners.

2. Integrability: our technique is design to smoothly become part ("plug into") general purpose, qualitative or semi-qualitative PRM frameworks, assuming that some threat categories will be treated qualitatively and across processes, while other process-related disclosure risks will be addressed quantitatively and in monetary terms).

3. Communicability: Our technique is targeted to improving consultation among process designers/owners in PRACTICE scenarios and other organizational roles on the alternative deployments of privacy controls developed within PRACTICE (see D21.1 - Deployment Models and Trust Analysis for Secure Computation Services and Applications). Improved understanding of mechanisms is a key factor for correct assessment of risk on the part of non-technical shareholders[39].

The details of integration between our technique and semi-qualitative risk assessment frameworks will be provided in a future deliverable. In order to make the present document self-contained, we now provide only a brief outline of the plug-in interface between a quantitative risk assessment technique and the outcome of a semi-qualitative PRM. Such an outcome is usually a *review of the likelihood and severity of threats, qualitatively assessed using an ordinal scale* of risk levels (Figure 4.4). In PRM threat analysis, the *severity* of a threat usually depends on the size and level (full or partial) of personal data that would be identified should the threat occur, while *likelihood* is estimated empirically by looking at the vulnerabilities of the supporting assets and the risk sources' capabilities of exploiting them.
Risk analysis is a crucial step of the *iterative PRM process* that typically involves (Fig. 4.5)

[2]The Office of the Privacy Commissioner of Canada has provided support for early-stage qualitative analysis of privacy risks in its PIPEDA Self-Assessment Tool, `http://www.priv.gc.ca/information/pub/ar-vr/pipeda_sa_tool_200807_e.cfm`.

[3]This approach was used in the US by the American Institute of Certified Public Accountants (AICPA) in their "Privacy Risk Assessment Tool"

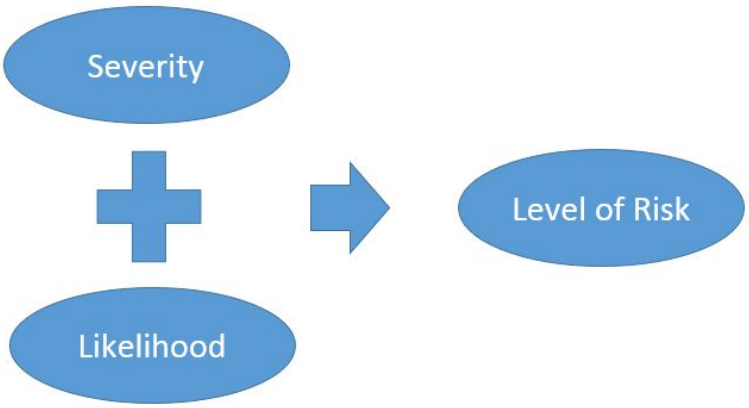

Figure 4.4: Severity/Likelihood Classification of Threats

1. *Establishing context* Study of the context where a business process P handling personal data will take place. For instance, in the e-health service context, process P could be the procedure used by chronically ill patients to request transportation services on vehicles with medical appliances (oxygen mask, defibrillator, life-signs monitors) needed for their illnesses.

2. *Privacy Risk Identification* Identification of feared events. For instance, the personal data of a patient together with her illness may fall in the hands of cyber-criminals.

3. *Risk Analysis* Severity/Likelihood analysis of specific threats that may bring those feared events about. In our example, the providers of an outsourced authentication service and the medical appliances reservation system could collude to link a patient's identity to the appliances she requested, in order to guess her diagnosis.

4. *Risk Treatment* The selection of countermeasures (privacy controls). In our example, a data scrambling component runs at a trusted site to randomly associate requests for appliances among patients sharing the same vehicle.

5. *Monitoring and Continuous Improvement* The revision of the business process P to incorporate execution of privacy controls, and the continuous monitoring of their operation.

6. *Communication and Consultation* The continuous communication to stakeholders of the improvement choices made on P, and the polling of their opinions.

Quantitative risk analysis discussed in the next section can be plugged in at the third step of the PRM framework of Figure 4.5, in order to provide monetary value of the risk related to certain threats considered crucial (in this case, collusion among participants to P to jointly reconstruct information that none of them would hold individually. This estimated risk value can be an important factor in estimating the expected Return Of Investment (ROI) of privacy enhancing technology.

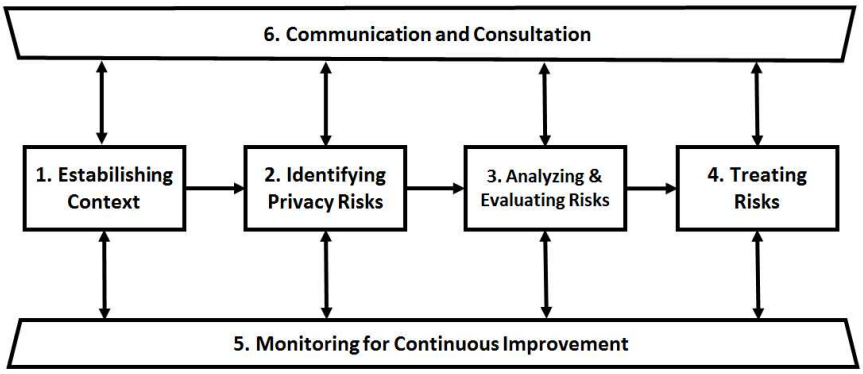

Figure 4.5: The structure of a PRM process

Chapter 5

A methodology for quantitative assessment of risks in cloud-based process execution

5.1 Threat Space: Disclosure Events

Any risk model must specify the threat event space where risks will be quantified. A well-known problem of applying general-purpose risk assessment frameworks based on Eq. 4.1) is the size of the threat assessment task, where each scenario taken into consideration introduces new families of threats. In an effort to be comprehensive, risk modelers have tried to to capture all threats, assets, vulnerabilities, and security concerns. Unfortunately trying to be exhaustive can put risk analysis requirements beyond the capabilities of both personnel and computing resources. Indeed, highly expensive assessment process was one of the pivotal reasons why early risk models failed to achieve widespread acceptance [62].
In this deliverable we chose focus on a single, albeit large, family of threats, namely *data process-related leakage threats*, i.e. *the disclosure of one or more information items to be exchanged in a multi-party protocol to participating parties who are not the originally intended recipients*. These threats are central to the cloud and in general to outsourced IT, and are due to many different factors. A major risk factor is potential collaboration among parties (including service and cloud providers) that put together partial information to reconstruct knowledge that was not available to them when taken individually. We remark that this unexpected information sharing is not due uniquely to the traditional notion of "collusion among rogue participants". Indeed, putting together information held by different parties may take place for different reasons, including (i) eDisclosure, i.e. the mandatory process of disclosing information to adversaries during litigation (See http://www.edisclosureinformation.co.uk) (ii) an information request from a regulatory authority[1] and (iii) inadvertent or dysfunctional behavior of employees. For the first factor, locating data and bringing it together within the timescales imposed by courts of law may generate leaks that are difficult to identify a priori even for experienced security auditors. This is due to the fact that processes and technological interfaces through which data is obtained for eDisclosure may differ to those that are usually followed for handling data on a day-to-day basis.
The second factor - the intervention of a regulatory authority -. is also difficult to evaluate a priori. For instance, e-mails containing bids for a auction held in one country may be stored

[1] As discussed in the first part of this deliverable, whether data is on premises or in the cloud, the obligation to comply with the demands of the court or regulatory authorities remains essentially the same.

on a server located in another jurisdiction, where a regulatory authority can ask the service provider - for reasons unrelated to the auction - full access to the storage of the mail server, without informing the auctioneer. This way, a third party would get to know in advance the outcome of the auction.
As one would expect, the third factor has the strongest documentary evidence. Findings from a recent global security study on data leakage, commissioned by Cisco and conducted by U.S.-based market research firm [16] revealed that data loss often results from insiders' dysfunctional behavior. The study polled more than 2000 employees and information technology professionals in 10 countries, including major EU markets. It selected countries based on their diverse social and business cultures, with the goal of better understanding whether these factors affect data leakage. The study's findings show that "insider threats" have the potential to cause greater financial losses than attacks that originate outside the company. Insider threat is traditionally characterized as an employee performing malicious behavior-through sabotage, stealing data or physical devices, or purposely leaking confidential information. Here, we focus on the threat deriving from partner organizations implementing sloppily an interchange protocol, or intentionally sharing information with other unauthorized parties. Organizations need to become aware that the insider threat is not just the rogue employee, but rather every partner (including suppliers of outsourced cloud-based services used by employees) and every device that stores information. For instance, a plaintext email containing a business offer sent in good faith through a "secure" cloud-based mail service poses a danger if disclosed by the cloud provider to a competitor of the original sender. Data leakage often results from this type of risky behavior by employees who are simply unaware that their actions are unsafe. Sometimes, this problem can be attributed to a lack (or inadequate communication) of corporate security policies (for instance "never send messages regarding company business via a public email service") to employees. Prevention of process-related data leakage is instrumental to preventing losses of intellectual capital exchanged in business process execution over the cloud. Also, it is receiving considerable attention for reasons of privacy protection of personally identifiable information. Information on customers, as well as their preferences, or even the size and timing of messages exchanged with them via outsourced services may be revealing information that should remain private at all times.
Summarizing, today it is very challenging even for experts in industry to identify, analyze and handle data leakage risks due to the complexity and diversity of business processes and of the underlying IT systems; the trend toward outsourcing and the cloud is further blurring the scenario . Many organizations have little visibility into where their confidential data is stored on the cloud or control over where that data is transferred during the execution of a process. Even when this insight is available, organizations lack a clear methodology to assess whether the process involves an acceptable level of risk. The methodology and models presented in the next Section are aimed at filling this gap.

5.2 Process Model

Let us now formalize our notion of business process model. [2]. 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 are used to generate messages to be exchanged during the business process' execution. Also, we denote by $\{I_{j,k}\}$ the impact of the disclosure

[2]It is important to remark once more that the aim of our model is enabling risk assessment; so the process representation is focused on risk-related rather than design-related aspects

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[3]. 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* [5] 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[4]
- *Local computations* $(A_i, f(), INFO_{i,ts})$ i.e. functions computed by actors on (portions of) locally held information at a given time.

Figure 5.1 shows a sample process:

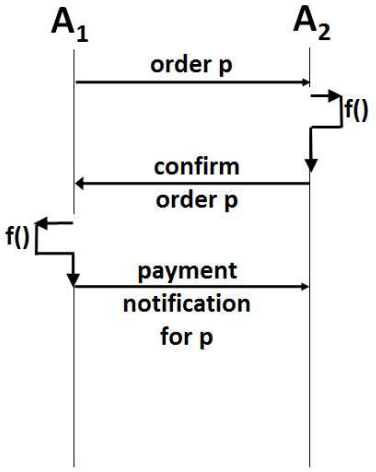

Figure 5.1: A sample process model

5.2.1 Process Model Assumptions

Our process model is completed by some additional assumptions. Here β denotes the probability of an event, however assessed (for our own economics-based probability assessment see Sect. 5.3):

- *Protocol efficacy*: Given a message delivery (A_s, A_d, m_{ts}), with $m_{ts} = INFO_s$ then $\beta_i(E_{sd}) = MAX$ for all actors A_i.
- *Information completeness*: Given a message delivery (A_s, A_d, m_{ts}), with $m_{ts} \leq INFO_s$, then $\beta_i(E_{sd}) = 0$
- *Strong local computation transparency*: Given a local computation $(A_i, f(), INFO_{i,ts})$, then $INFO_i = INFO_{i,ts} \cup f(INFO_{i,ts} \cup S_f)$ for $t \geq ts$, where S_f is the *specification* of f as an algorithm or a closed formula.

[3]Of course, the impact of disclosing an empty item is always 0.
[4]For the sake of simplicity, in our model we assume synchronous clocks and instant message delivery.

input1	input2	output	garbled computation
k_q^0	k_p^0	k_z^0	$E_{k_q^0}(E_{k_p^0}(k_z^0))$
k_q^0	k_p^1	k_z^0	$E_{k_q^0}(E_{k_p^1}(k_z^0))$
k_q^1	k_p^0	k_z^0	$E_{k_q^1}(E_{k_p^0}(k_z^0))$
k_q^1	k_p^1	k_z^1	$E_{k_q^1}(E_{k_p^1}(k_z^1))$

Figure 5.2: Garbled computation for an AND gate

- *Belief propagation*: Given a message delivery (A_s, A_d, m_{ts}), then for $t \geq ts$, $\beta_i(E_{sk}) = \beta_i(C(A_d, A_k))$ for $k \neq d$, where $C(A_d, A_k)$ denotes the event of information sharing between A_d and A_k.

It is important to remark that the *local computation transparency* assumption implies that any party computing a function $f()$ over its local data becomes aware of the results of that function as well as of its *specification* S_f, represented e.g. as a computer program. However, research has shown that this assumption may be weakened by *obfuscation* or *garbling* techniques [6].

5.2.2 Garbling Outsourcing Scheme

Garbled circuits, a classical idea rooted in early work by Andrew Yao, are a well-known example of obfuscation techniques. Here, we follow the literature [6] to briefly describe a *garbling outsourcing scheme* corresponding to Yao's garbling technique. The purpose of our simplified description is making this deliverable self-contained by showing how *obfuscation is represented within our process model*. The next release of this deliverable will show how to represent schemes made available by other PRACTICE work packages.
Let us assume Alice wants Bob to compute for her a function $f()$ on a set of inputs, some of which are held by herself and others by Bob, without sharing with Bob the function specification S_f . At a high level of abstraction, the scheme works as follows: Alice creates a "garbled circuit", i.e. the specification S'_f of a garbled function $f'()$ having the same input-output table as $f()$, and sends it to Bob. Bob uses S'_f to build $f'()$, compute it with his inputs B and returns the result to Alice. The result of $f'(B, x)$ evaluation with $x = A$ (where A is Alice's inputs) coincides with the function $f()$ that Alice wanted Bob to compute; but by computing $f'()$, Bob has learnt nothing about S_f. Note that in this scheme Alice does not send her inputs to Bob; rather, her inputs are encoded into the "garbled circuit" in such a way that Bob can not determine what they are. As an example, assume that Bob has $x = 2$ bits, (a, b), and Alice has $y = 2$ bits, (c, d). The function $f()$ is:

$$f(x, y) = (a + c) \vee (b + d) \tag{5.1}$$

For the construction of the garbled circuit, one simply constructs a new truth table for each gate in the original circuit. A sample truth table for an AND gate is shown below (Table 5.2), with inputs p, q and output z. Alice picks two random keys for each wire and obtains the garbled truth table by encrypting the output-wire key with the corresponding pair of input-wire keys
After Bob has received the garbled specification $S_{f'}$ and the corresponding truth tables, he still needs Alice's inputs before he can evaluate the function. Bob can get these inputs using a 1-out-of-2 instantiation of Rabin's *oblivious transfer* protocol [5]

[5]In an oblivious transfer protocol, a sender transfers one of potentially many pieces of information to a receiver, but remains oblivious as to what piece (if any) has actually been transferred [49].

Once Bob has received the input values from Alice via the oblivious transfer protocol, he can "decrypt" each of the gates, and using his own inputs he can evaluate the circuit. Today, efficient garbling schemes are available achieving privacy as well as obliviousness and authenticity, the latter properties being needed for private and verifiable outsourcing of computation. Highly efficient block-cipher-based instantiations of garbling schemes have been described in the literature.
For our purposes, it is sufficient to observe that when a garbling outsourcing scheme is in force within a process, a weaker assumption (*weak local computation transparency*) can be adopted for our business process model, where the party executing a local computation $f()$ learns the output of the function, but not its specification.
More formally, let us consider a process P including a local computation $(A_i, f(), INFO_{i,ts})$. Let $G()$ be a functional acting on the $f()$ function specification S_f, so that

$$G(S_f) = S_{f'} \tag{5.2}$$

We call $G(S_f) = S_{f'}$ a *garbled specification* of $f()$ if and only if $f'(x) = f(x)$ for all inputs x and:

$$t \geq ts : (A_i, f(), INFO_{i,ts}) \rightarrow INFO_i = INFO_{i,ts} \cup f'(INFO_{i,ts}) \tag{5.3}$$

It is important to remark that the computation of $G(S_f) = S'_f$ *can itself be a local computation of the process* P. This way, any actor can outsource a local computation to another actor, who will compute the garbled function under our weak transparency assumption.

5.3 Probability Model

A key activity when performing a quantitative risk assessment is the estimation of the uncertainty present in a variable. The content of a memory cell, for instance, seen as a binary number, has an average, a maximum and a minimum value. Therefore, to fully describe its behavior one needs a function, mapping binary numbers to their respective probabilities. If the size of memory cell is 1 byte, this function (called *Probability Mass Function* (PMF)) would map each integer between 0 and 255 to a probability value between 0 and 1. In the case of a continuos variable, it is not meaningful to talk about a probability for a given specific value, since there is an infinite number of values between the minimum and the maximum values, and therefore (given the classical definition of probability) the probability of each individual value is equal to zero. Instead, when describing the uncertainty of a continuous variable, like the height of a human, one uses the concept of a *Probability Density Function* (PDF) that maps intervals of possible values to probabilities. In this Section, we will concentrate on the discrete case in this work, since it can be readily applied to discrete events like threats. A major problem for the practical application of our risk assessment methodology is that the needed PMFs are not readily available, and have to be derived from available information and knowledge. Several methods have been proposed in the literature for the derivation of PMFs and PDFs in risk assessment. The choice of an appropriate method depends on what information and knowledge is available. In the remainder of the Section we will discuss how to build a PMF associating subsets of actors in 2^A to the probability of them *colluding*, i.e. putting together the information they know. As we will see, this PMF will be computed in two steps. First, we will deriving the PDF associated to a continuous variable, the degree ϕ of perceived unfairness in a business process resource allocation. Then, we will use the probability values of the PDF to compute

a discrete PMF, associating to each subset of actors the probability that the subset members will decide to put together the information they hold.

5.3.1 Estimating the probability of Collusion

Any multi-party business process is exposed to the possibility that a subset of the members starts behaving in an unprescribed way, sometimes with the hidden purpose of damaging the process. Information leakage events from individuals in collaborative processes involving information processing and exchange or collusion among multiple actors are often originated by this kind of contingency.
The schema for the computation of the probability of collusion among actors is the following. In order for a collusion between two actors to take place we consider that some enablers are needed *motivation*, *opportunity* and *means*.
In our model the *means* are already made available to an actor by virtue of the actor *role*: e.g. the capability of accessing some information present in some cloud infrastructure is granted to the administrator of that infrastructure by its own role. Thus the definition of the actor role already contains the means for the type of violation the actors can perform and are already accounted for by the process model.
Motivation and opportunity are the remaining two factors to be accounted for: they can be treated probabilistically.

- *Motivation* is an individual actor property. The presence and strength of *motivation* to defect by an actor i will be accounted for by a factor ϕ_i, modeling the degree of insatisfaction by the actor. The probability that the presence of motivation translates into an actual violation is modeled by a function $p_i^c(\phi_i)$, where c indicate the context; when the function is the same for all the actors one can indicate it by $p^c(\phi_i)$. As we will see seen in the next subsection this probability can be obtained based on the elicitation of experts' opinion.

- *Opportunity* (of interaction) expresses the event that two or more actors interact. It is a relational property: it can involve two or more actors. The probability that two or more actors interact can be expressed by a factor $Q(S)$ where S is the actor subset. Also this probability can be obtained based on the elicitation of experts' opinion. The probability can be modeled by different levels of complexity and depends on the process model.

The role of motivation and opportunity in the computation can be exemplified as follows: if actor 1 with and actor 2 with motivation ϕ_1 and ϕ_2 and with consequent individual probability of violation $p^c(\phi_1)$ and $p^c(\phi_2)$ interact with probability $q(\{1,2\})$ they will collude with probability $p^c(\phi_1)Q(\{1,2\})p^c(\phi_2)$. In this case we assumed that the probabilities of defection $p^c(\phi_i)$ are conditionally independent, i.e. that their values are independent given the context. This assumption can be made in many cases of practical interest.
Now we discuss first the derivation of the probability of defection, based on the concept of Shapley Value, then the discuss the modeling of the probability of interaction; finally we will provide an illustrative example of computation.

5.3.1.1 From the Shapley Value to the probability of defection

At the basis of dysfunctional behavior by actors in a business process is often the *unfairness of the redistribution of payoffs* in the process: for example, a benefit allocation structure that

responds to organization efficiency more than to fairness requirements. Indeed, a process configuration yielding the highest total surplus does not necessarily guarantee a fair distribution of this surplus. Efficiency says nothing about equity or fairness, i.e. *distributive justice*.
The problem of how profits of a coalition should be redistributed is a well-known one (it is an instance of the general problem of distributive justice). There are several kinds of solutions to the problem. Due to the subjectivity of satisfaction criteria for each agent an objectively optimal solution cannot in general be attained, however a solution fulfilling some largely accepted requirement can be obtained by following the prescription dictated by the so called *Shapley Value*[4]:
"given a coalition, the contributions of the actors to the process, and the value of the surplus value produced by the process, the Shapley value yields a unique ideal allocation of that value fulfilling some largely accepted requirements. With N actors, this solution can be visualized as a point on a hyperplane in an N-dimensional space."
The Shapley approach can be applied to the actors of an organization to find the fairness point and compare it to the point representing the current allocation of the value in an organization. In general, the distance between these two points provides for each actor an estimate of its individual probability of defection: the closer the two points, the more likely there will be no dysfunctional behavior on the part of that actor. Should the convenience become too low, the actor will be tempted to behave un-cooperatively. Such behavior may damage to the overall business process.
We provide now a formal definition of the Shapley value. The idea behind it is that each player should be given a payoff equal to the average of the contribution that he makes to each of the possible coalitions.

Let us consider a general game with a set $\mathcal{N}$ of N participants. The Shapley Value is defined as an allocation of payoffs: a payoff u_i for each actor $i \in \mathcal{N}$. Any subset of players in N is a potential *coalition* C. A coalition can strike deals among its own members to exploit all the available knowledge for mutual advantage. Combinatorially, there are $(2^N - 1)$ possible coalitions altogether, including the so-called *grand coalition* consisting in $\mathcal{N}$ itself (and disregarding the empty set).

It is usual to call *security level* of a coalition C the quantity $s(C)$ expressing the total surplus that its members can achieve on their own even if the non-members took the action that was the worst from C's perspective. An allocation $(x_1, x_2, \ldots, x_N)$ is a list of amounts for the players (they are shares of a total value, and add up to the total added value), and it is said to be feasible if allowed by the rules of the game. A feasible allocation is blocked (i.e. not even considered) by a coalition C if $a(C) > \sum_{i \in C} x_i$, i.e. if the allocation values add up to an amount which is less the security level of the coalition.
We call *core* the set of allocations that cannot be blocked by any coalition: it contains all possible reasonable deals. The core can be a point, a range or a general set. For some games, it can even be empty. However, the core has some desirable properties. Since it cannot be reduced further by any groups searching for a better deal, including the grand coalition, it can be shown that it is *Pareto efficient*, i.e. that no allocation outside the core will improve everyone's payoff simultaneously. Note that even the core, being defined on the base of an inequality over a sum of the allocation array, does not give any guarantee over the distributive justice of an allocation: the elements of the core will all represent efficient allocations, but some will be fairer than others. In order to produce each coalition one has to run ideally over all the permutations of actors: each ideal ordering of the actors corresponds to a non-decreasing

surplus value achieved by the members up to the considered index, when arriving at the actor i, whose Shapley Value is being computed, one has to take note of the added value introduced him, the Shapley Value for an actor is then given by the average over all permutations of those added values, which we denote by Δ_i:

$$u_i = \frac{1}{N!} \sum_{\pi} \Delta_i(\pi) \tag{5.4}$$

where the index π runs over all the permutations of N objects.
Eq. 5.4 can be rewritten in a more computable form by taking into account that, when scanning a permutation the actors following i (the trailing actors) are irrelevant to the computation of that actor's added value, and that, to the same computation, the order in which the actors preceding i (the leading actors) are ordered, is irrelevant, provided that the composition of the set of those actors is the same.Denoting by C such a set, the payoff u_i defined above can be computed as follows:

$$u_i = \sum_{C \subseteq \{\mathcal{N} \setminus i\}} \left[\frac{N!}{(|C|)!(N - |C| - 1)!} \right]^{-1} \left(\; s\left(\{C \cup i\}\right) - s(C) \; \right) \tag{5.5}$$

Here C represents the set to which i brings his contribution as an additional actor (the leading set), $|C|$ the cardinality of such a set, and N the overall number of actors, while $\Delta_i(C) \equiv \left(s\left(\{C \cup i\}\right) - s(C)\right)$ is the difference in security levels. The quantity in square brackets is a combinatorial factor which accounts for the fact that all the permutations have the same probability $1/N!$, that for a given leading set C there are $(|C|)!$ equivalent orderings and that for the trailing set, consisting in $(N - |C| - 1)$ elements there are $(N - |C| - 1)!$ equivalent orderings.

From the Shapley value to adverse event probability Given the Shapley value for each actor in a subset, we first define a PDF (to be used in the computation of the collusion PMF) as follows. Let $\delta_i = u_i - x_i$ be the difference between the Shapley Value and the actual resource allocation for that actor, i.e. the benefit the actor expects from taking part to the process. If this difference is positive, it means that the actor is under-rewarded for his contribution and there may be a positive probability that this causes a defection; if, instead, it is negative, the actor is over-rewarded and this discrepancy will not contribute to its probability of defection; one needs also to relate the discrepancy, when positive, to the absolute value of u_i.

For all the above considerations the motivation factor ϕ_i for an actor i can be defined as follows

$$\phi_i \equiv \frac{\theta(u_i - x_i)}{u_i} \tag{5.6}$$

where $\theta(.)$, is a filter function defined for an argument $z \in R$ as

$$\theta(z) \equiv \begin{cases} 0, & if \quad z < 0 \\ z, & otherwise \end{cases} \tag{5.7}$$

As a result ϕ_i is a non-negative quantity. Notice also that ϕ_i is a decreasing function of x_i and that it achieves it maximum at $x_i = 0$ where $\phi_i = 1$. Overall $\phi_i \in [0, 1]$.

Elicitation of expert opinions. By itself, the above defined ϕ factor does not allow to compute the PMF of each subset of actors colluding to share the information they hold. The factor has to be first translated into a PDF by a function $p^c(\cdot)$ which takes into account not only obvious constraints – namely $p^c(\phi = 0) = 0$, and the non-decreasing nature of the function– but also contextual conditions, represented by the exponent (c), often difficult to model mathematically.
We derive such a PDF based on expert evaluation of the contextual elements. This is an instance of the well-known problem of eliciting experts' knowledge in order to obtain PDFs for a continuos variable, given the value of different context parameters. The result of this type of analysis is typically a PDF over the value of one or more variables. In our case, we determine a probability distribution over the values of ϕ.
PDFs fall into two categories: *non-parametric* and *parametric* Distributions Functions (DFs). A parametric distribution is based on a mathematical function whose shape and range is determined by one or more distribution parameters. These parameters often have no obvious or intuitive relationship to the shapes that they define. An example of these are the Beta density or the Weibull distributions. Non-parametric distributions on the other hand, have their shape and range determined by their parameters directly, in an obvious and intuitive way. Their distribution function is a mathematical description of their shape. Typical non-parametric distributions are Triangle, Bernoulli and Discrete.
Despite the positive qualities of parametric PDFs, there are cases - like ours - where the available information on studied quantity simply does not allow the selection of a parametric PDF. In the case of ϕ, the number of classes (i.e., the levels of perceived unfairness) and the variance of the elements of each class is small, and the modes of the individual classes are far apart. Then the DF will exhibit more than one mode. Because of the unimodal character of the commonly used PDF, this situation requires the use of a nonparametric or empirical DF. The name nonparametric is actually somewhat misleading as these functions, in fact, do have parameters. However, as opposed to parametric DFs, whose parameters represent important statistics of the observed data, the nonparametric PDFs use the entire data set as parameters. A well-known, simple example of nonparametric PDF is a *frequency histogram*, i.e. a bar chart of values which provides a frequency-based approximation of the PDF. It can be derived from the ordinary value histogram simply by re-scaling bars to make the integral equal to unity.
In our methodology we use a non-frequency-based solution, the Bézier Distribution Family based on Bézier curves. Bézier curves have often been used in computer graphics to approximate a smooth (continuously differentiable) function on a bounded interval by forcing the Bézier curve to pass in the vicinity of selected control points in two-dimensional Euclidean space. Given a continuous random variable whose space is a close bounded interval, then in principle we can approximate its distribution function arbitrarily closely using a Bézier curve by taking a sufficient number of control points with appropriate values for the coordinates. Hence, the computation of the probability of adverse events involved the following steps:

- Expert opinion is elicited to determine the probability density as a function of the percentage Shapley value deviation. Bézier curve based densities are suitable candidates for representing expert opinions;
- Then, given a specific instance of the collaborative process definition, we compute the numerical value of $p_i^c(\phi_i)$ for every actor

Expert opinion is thus used to translate contextual factors into the shape of the function $p^c(\cdot)$. Then, the PMF of collusions within subsets of actors can be computed by assigning

higher probability of collusion to subsets composed of actors with the highest and possibly homogeneous ϕ value.

5.3.1.2 The probability of interaction

The simplest way of modeling the probability of interaction, is by the use of a simple combinatorial factor counting the number of pairs. This is equivalent to adopting the following two approximation assumptions

1. (*Uniformity approximation*) the probability of interaction is uniform over all the actor pairs:
$$Q(\{i,j\}) = q \quad \forall i,j \in A$$
2. (*Linked pair approximation*) the probability of interaction of triplets of actors, quadruplets of actors, and so on, is completely accounted for by actor-pair interactions: e.g.
$$Q\left(\{1,2,3\}\right) = Q(\{1,2\})\,Q(\{2,3\})\,Q(\{3,1\})$$
in other words higher order interaction can be expressed in terms of linked-pair second-order interaction.

If expert opinion endorses this approximation and provides the value of q, then the probability $P^c(S)$ that in a given context c a set S – consisting of a number $|S|$ of actors – collude can be computed as a product of the individual motivation related factors and the opportunity related factor as follows: the former is expressed by a product of the $p^c(\phi_i)$, the latter by q to a power corresponding to the number of pairs in the set $r = |S|\,(|S|-1)/2$

$$P^c(S) = q^r \prod_{i \in S} p^c(\phi_i)$$

A slightly more complex and in principle more accurate representation is obtained by lifting assumption 1) above: in that case the probability of interaction is no longer uniform over the pairs. Expert opinion can be elicited to obtain the values $Q(\{i,j\}) = q_{ij}$ specific of each pair, which is likely to depend on the structure of the process. Then the probability $P_c(S)$ can be computed as

$$P^c(S) = \prod_{i,j \in S} q_{ij} \prod_{i \in S} p^c(\phi_i)$$

Lifting also the *linked-pair approximation* assumption 2) can provide an even more accurate estimate of the probability of collusion at the price of a further modeling and elicitation effort: in the extreme case expert opinion will have to provide the full *structure function* $Q(S)$ for all $S \in 2^A$. The probability $P_c(S)$ will be computed as

$$P^c(S) = Q(S) \prod_{i \in S} p^c(\phi_i)$$

The computation of $P^c(S)$ under either set of assumptions can be computationally and operationally expensive (expert opinion elicitation may be costly), however the procedure can be made more efficient by giving higher priority to the modeling of the most likely events: one can consider first the actors with the highest values of ϕ, and in a non-uniform model the pairs with the highest values of $q_i j$ and so on.

5.3.1.3 Example

At this point, we can compute for each subset in 2^A the probability of collusion, and subsequently the impact of such a behavior, as explained in the next subsection, so that finally the corresponding individual actor risk – defined by the product of the probability of unilateral attack and the impact of the unilateral attack – can be computed.
A sample computation of the probability of collusion of three actors participating into a secure computation is given here, a graphical representation of the result is given in Figure 5.3. Let us assume that in a secure computation are involved the following actors $\{IN, COMP, RES\}$, in charge respectively of the input, computation and output phases, plus a fourth actor ACC in charge for the accounting aspects of the business. The total revenue from the computation is $s = 400$. A simple computation of the Shapley Value of each actor yields, due to the irreplaceability of the actors we have

$$u_{IN} = u_{COMP} = u_{RES} = u_{ACC} = 100$$

The actual payoffs they obtain from the computation are however

$$x_{IN} = 80 \qquad x_{COMP} = 70 \qquad x_{OUT} = 90 \qquad x_{ACC} = 160$$

Their motivation factors are respectively

$$\phi_{IN} = 0.8 \qquad \phi_{COMP} = 0.7 \qquad \phi_{OUT} = 0.9 \qquad \phi_{ACC} = 0$$

Indeed the actor ACC has no motivation for defecting this advantageous deal.
The elicitation of expert opinion has determined that the probability of defection as a function of the motivation factor is quadratic and is given by $p^c(\phi) = \phi^2$ (the function is defined in the interval $[0, 1]$). As a consequence the individual probabilities of defection are

$$p^c_{IN} = 0.64 \qquad p^c_{COMP} = 0.49 \qquad p^c_{OUT} = 0.81 \qquad p^c_{ACC} = 0$$

We exclude the actor ACC from the following discussion, since it is excluded that she will participate to any coalition. This concludes the individual actor level analysis and provides the top level of the representation in Figure 5.3.
Now we focus on the opportunity of interaction an on the determination of the factors q_{ij}. We observe the process description and note that the pair $\{IN, COMP\}$ and the pair $\{COMP, RES\}$ are prescribed to interact by the process itself, hence we set the probabilities of interaction $q_{\{IN,COMP\}}$ and $q_{\{COMP,RES\}}$ to 1. The elicitation of expert opinion determined that the probability that IN and RES get in contact is $q_{\{IN,RES\}} = 0.3$. The resulting level-pair estimate of the probability of collusion is therefore given by

$$\begin{aligned} p_{\{IN,COMP\}} &= p_{IN} \times 1 \times p_{COMP} &&= 0.3136 \\ p_{\{COMP,RES\}} &= p_{COMP} \times 1 \times p_{RES} &&= 0.3969 \\ p_{\{IN,RES\}} &= p_{IN} \times q_{\{IN,RES\}} \times p_{COMP} &&= 0.1555 \end{aligned}$$

The expert opinion has granted also that in this case one can apply the linked pair approximation. The probability of the grand-coalition of the three actors therefore has probability

$$\begin{aligned} p_{\{IN,COMP,OUT\}} &= p_{IN} \times 1 \times p_{COMP} \times 1 \times p_{RES} \times q_{\{IN,RES\}} \\ &= 0.64 \times 0.49 \times 0.81 \times 0.3 \\ &= 0.0762 \end{aligned}$$

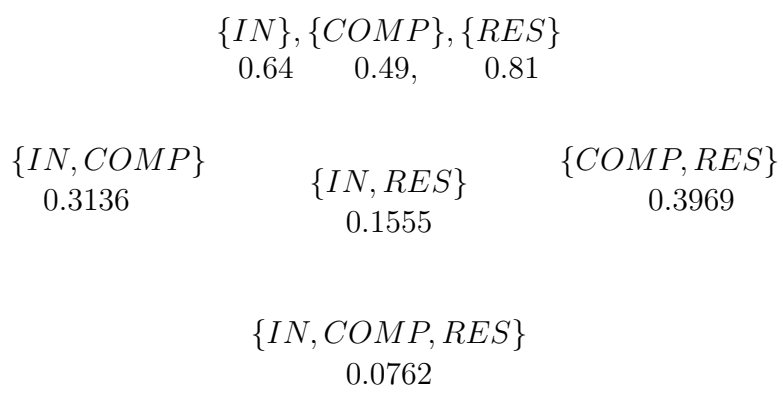

Figure 5.3: The outcome of the computation of the probability of collusion for the actor set

The overall picture is shown in Figure 5.3. The probabilities associated to singletons express their likelihood to defect, while the ones associated to subsets express the likelihood of the group to collude and exchange informations
This example concludes the discussion of the methods for determining the probability of dysfunctional behavior, which in the case interesting to us is limited to the collusion within subsets of actors to share the information they hold. We move now to the calculation of the impact in some representative case.

5.4 Impact Assessment

The technique we use for estimating impact of information disclosure is loosely related to the one we just used for our probability estimate. Let us start with an example: if an information item sent via email via a cloud-based mail service contains, say, an attachment with the design information of a new product, what will be the impact of its disclosure? To answer this question, we can use two different approaches:

1. perform an accurate analysis to precisely quantify the impact of the disclosure, e.g. in terms of the financial loss the company would occur in when a new competitor enters the market (an event that would certainly happen once the information item containing the new product design has been disclosed to current competitors)

2. use an arbitrary discrete unit and quantify the perceived impact of loosing such a message to a conventionally high level, corresponding to a perceived "disaster".

In terms of our methodology, our process-oriented way to perform impact analysis relies on quantifying the *Value of Information* (VoI) for each knowledge set K_S potentially reconstructed by a subset $S \in 2^A$ of the process actors.

5.4.1 Value of Information Analysis

VoI has been defined as the *analytic framework used to establish the value of acquiring additional information to solve a decision problem.* In the risk management domain, VoI has been successfully used since the Sixties in several areas of research including engineering and environmental risk analysis [28].
From a purely rational perspective, it is clear acquiring extra information is only useful for an actor A if knowing it has a significant probability of modifying its behavior. Classic VoI

analysis typically involves constructing a complex decision-analytic model to fully characterize all information items available to each process actor, the loss each actor would incur should these items become known to other actors, the costs of interventions that could be executed to prevent them. This comprehensive approach to VoI often turns out to be prohibitively expensive for use in prioritizing interventions [27]. As alternatives to full VoI, we identified three approaches to analyzing the value of information that are less burdensome:

1. The *conceptual* approach to VoI, where context information is used to provide informative bounds on the value of information without formally quantifying it through modeling. For instance, the VoI of the design information about a device that is already available on the market cannot be higher than the cost of reverse-engineering the device itself;

2. The *minimal* approach to VoI, which is possible when evidence of the net benefit of holding a piece of information, are readily available from existing research. For example, the VoI of the design information about a device that is currently available on the market cannot be higher than the net profit coming from its sales to its current supplier.

3. The *maximal* modeling approach to VoI, where the value of an information item is estimated from previous VoI studies concerning similar information in different contexts. For instance, the VoI of the design information about a solid-state storage device is quantified according to previous VoI studies on disks.

These three low-cost VoI methods create the possibility for VoI analysis to be applied in priority-setting of risk-alleviation countermeasure, and raises the question about how the use of VoI to assess disclosure risk in the framework of our methodology.

Our methodology takes a process-oriented view of VoI, in order to assess the impact of information disclosure. Let us consider once again a set of actors $A = \{A_1, \ldots, A_n\}$ who take part to a business process P, and the expected benefit for each actor A_k, Ben_{A_k} resulting from the execution of P. The starting point of our VoI analysis of P is to consider the *Value of Total Information* (VoTI), i.e. answering the question "What would be the change to Ben_{A_i} should A_i know all information (local memory plus messages) held by the other actors of P?". If there is no such change, then achieving extra information is worthless. If such a change exists, then the impact on A_k of A_i's ($i \neq k$) complete knowledge can be estimated as the corresponding change in the value of Ben_{A_k}.
For the security-aware process designer, our simple VoTI provides a useful upper bound, because it tells the maximum value that any information held by other actors may have for each participant to P. If that value is negligible, or achieving that information would cost more than that, a rational actor will not pursue disclosure any further (i.e., it would not enter agreements for information sharing with other actors).
A different type of check involves looking at the *Value of Partial Information* (VoPI). For any process participant A_i, getting to know some information beyond the one that is strictly necessary to carry out its part in the process (e.g., the messages exchanged among other actors, or the content of another actor's local memory) may or may not bring a benefit, i.e. a change in Ben_{A_i}. For each subset K of knowledge items used in the process, VoPI focuses on (i) checking whether the benefit of knowing K would match the cost of collecting it and (ii) quantifying the impact of each actor A_i getting to know K on the benefits Ben_{A_k} of the other participants (for $i \neq k$) .

5.5 Methodology

Managing risks related to the execution of a business process P in presence of threats is itself a process (usually called *risk management process*, in symbols $M_{R(P)}$) where alternative techniques for dealing with threats are compared. The output of $M_{R(P)}$ is a *risk alleviation strategy*, which consists of modifications to P that have some effect on the risk of executing it, including the introduction or removal of security controls like the ones defined within PRACTICE. In this Section, we put forward a methodology for comparing alternative risk alleviation strategies. Our methodology does not provide specific guidance on the choice of mechanisms that will actually counter the threats; rather, it allows comparing the residual risk of competing risk strategies. Although qualitative comparison is supported, the methodology aims to quantitative cost-benefit calculations, assessments of risk tolerance, and quantification of preferences involved in $M_{R(P)}$.

Before describing the methodology we remark that *methodological choices* made in probability assessment strongly affect what can be practically done within $M_{R(P)}$. For instance, it is sometimes possible to ask the users for a rough estimate of perceived probabilities and impacts and then multiply them to get *risk coefficients*. Such coefficients can be used for a "quick-and-dirty" comparison of versions of P that include different security controls. On the other hand, a (more costly) best-effort computation of probabilities and impacts allows to use quantitative estimates of $R(A, E)$ to drive the organization's choice between alternative implementations of P, by comparing $R(A, E)$ values to the cost of adoption/deployment of proposed security patches or controls.

We are now ready to provide a high-level description of our risk analysis methodology. Figure 5.4 shows the flowchart.

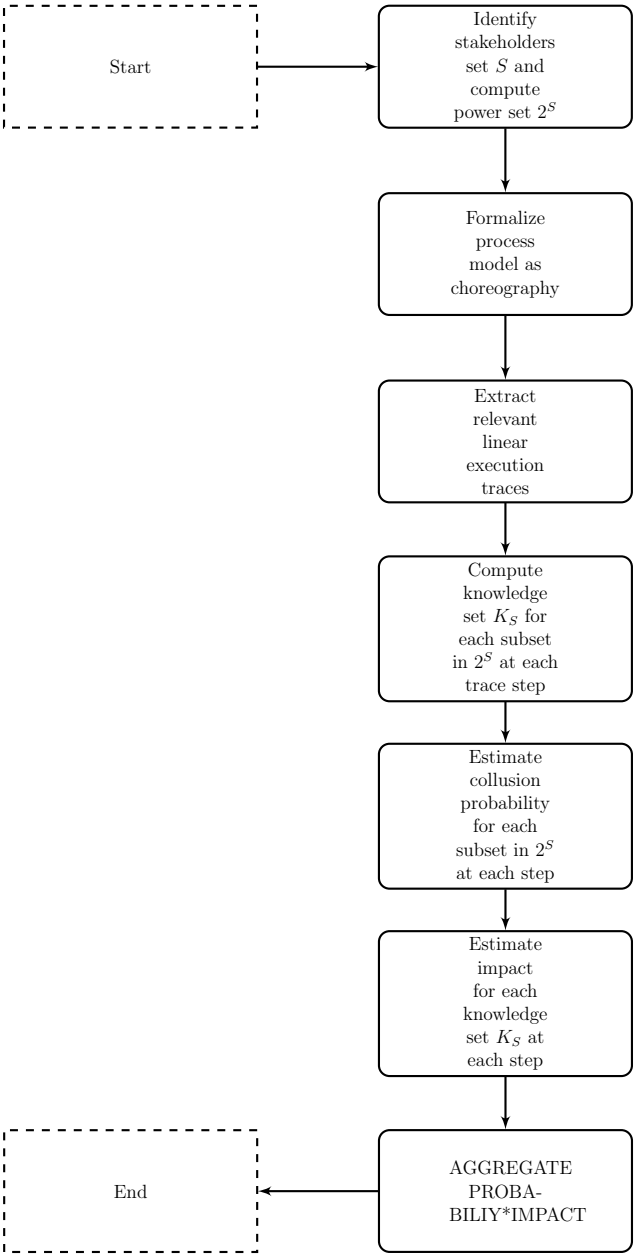

Figure 5.4: Our methodology's flowchart

Let us now go through our risk assessment methodology step by step:

- The first step is the *stakeholder identification*, where we identify the actor set A of our business process P and compute its power set 2^A. In our approach, process stakeholders include all participants to P. Namely, our actor set includes *all actors who, according to the risk assessor, may in any way get the capability of reading (or writing) information shared during P's execution.* While this first version of the methodology does not deal with the representation of individual actors', we plan to use state-of-the techniques for estimating the *salience* of actors' attributes with respect to probability assessment [42], [55] in the next version.

  However, as we shall see in the following Sections (Sect. 5.6.1), actors in A can be further refined by type according to their role in the computation.

- The second step consists in the *formalization of the business process model*, using the syn-

tax introduced in Sect.5.2, which represents two types of actions: (i) *message exchanges* and (ii) *local computations*. It is important to remark that while execution-oriented process models usually contain control structures like conditions and loops [56], our process model syntax expresses all possible execution paths *independently*, i.e. as separate models. The next step takes care of this.

- The third step consists of *process streamlining*, which includes *loop unrolling* and *re-encoding of conditions as parallel paths*. Here we do not enter into the details of business process streamlining, as process improvement techniques have been deeply studied since the Eighties and are discussed in detail in the technical literature (see for instance the rich bibliography of [56]). However, software toolkits supporting our methodology will have to provide guidance w.r.t. process streamlining.

- The fourth step, *identifying reconstructible knowledge*, consists in computing the *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*.

- The fifth step consists in *estimating the collusion probability* for each subset S in 2^A at each step of the process P. 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 (Sect. 5.1)

- The sixth step consists in *estimating the disclosure impact* of K_S for each subset S in 2^A at each step of the business process P

- The seventh and final step consists in *aggregating the products* 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.

It is important to remark once more that the choice of the probability assessment method will strongly depend on the threats whose probability we are trying to assess. When the threat is the dysfunctional behavior of one or more of the stakeholders of a business process, like putting in common the knowledge they hold at a given time, alternatives for probability assessment go from our analysis of the micro-economics model below the process (see Sect.5.3) to the assessment of the perceived level of the event's likeliness in terms of the social network of relations between the individuals involved in the process[6]. In turn, the *impact quantification method* used to assess disclosure impact can go from a simple ordinal prioritization of levels of information sensitiveness to complex analysis of potential loss that would be caused by the disclosure of specific data items. In the next Sections we will see some specific assessments targeted to cloud processes.

5.6 A Cloud-based Process Model

In this section we specialize our process model presented in Section 5.2) to describe cloud-based computations, using Bogdanov et al.'s representation of actors[7].

[6] As an alternative, a *Bayesian Belief Network* (BBN) [63] can be used to model the process, by taking into account its different actors and their mutual influences. The BBN can be exploited in different ways, especially to support identification and evaluation of risk control options at the organizational level.

5.6.1 The Cloud Actor Set

We enrich our representation of our multi-party business process actors and make it suitable for representing cloud-based computations. To this end, our actor set A becomes (non-necessarily disjoint) a 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).

5.6.2 The Cloud Process Model

The following constraints - looser versions of the ones in [7] - are in place for our cloud model:

- *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 5.5 shows a sample visual representation of a cloud-based process, where a buyer sends messages to two sellers who respond with their offers:

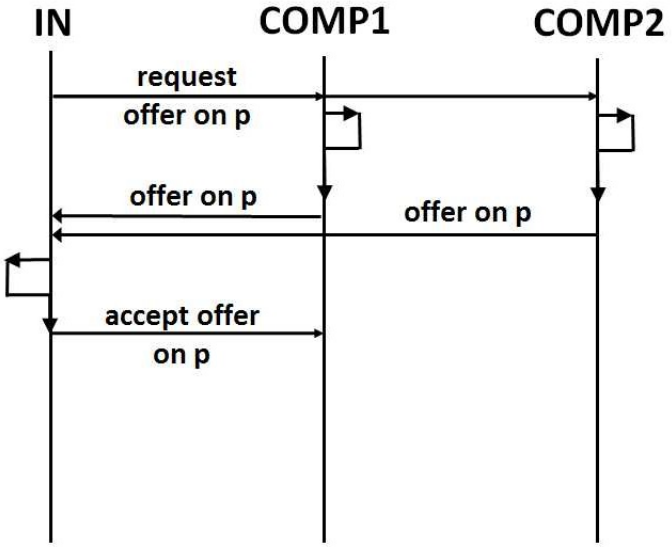

Figure 5.5: Visual representation of a sample cloud process model

5.7 Impact and Probability Assessment

For each subset $S \in 2^A$, we can now compute the risk of disclosure for information shared within 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:

$$I_{S,k,t} = \sum_{p=1}^{h} I_{j_p,k} \tag{5.8}$$

and, in words, as the damage that members of S can do to A_k by getting to know 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. This risk can be written as follows:

$$R(A_k, E_S) = P_{S,t} I_{S,k,t} \tag{5.9}$$

Assuming that collusions happen independently, we can also write the total risk for A_k taking part to the process, as follows:

$$R(A_k, 2^A, \infty) = \sum_{S \in 2^A} I_{S,k,\infty} P_S \tag{5.10}$$

However, it is clear that information sharing events - collusions - are in general not independent. We shall take care of dependency in the next version of our methodology.

5.7.1 Sample Assessments

Let us start with a very simple example: a business process where a client uses an outsourced computation service to add two integer numbers and another one to publish the result. In this case, we have the actor set $A = (IN1, COMP1, RES)$ where actor $IN1$ holds the information item $INFO_1$ containing the two summands $INFO_1[1]$ and $INFO_1[2]$, actor $COMP1$ is the outsourced services that computes the addition, while actor RES publishes the result. The process is represented by the choreography shown in Fig. 5.6, where the input actor $IN1$ sends $INFO_1$ to $COMP1$, who computes the desired local function $f(INFO_1) = INFO_1[1] + INFO_1[2]$, i.e. adds the two summands and sends the result to the result node RES who outputs it.
According to the definitions given in the previous Section, the (possibly empty) *common knowledge* of any subset of actors $S \in 2^A$ at time t, namely $K_S(t)$, is composed of the $INFO$ items that have been received in their entirety by all members of S at or before time t. The power set 2^A of the actor set is the simple Boolean lattice:

$$\{IN\}, \{COMP\}, \{RES\}$$

$$\{IN, COMP\} \qquad \{IN, RES\} \qquad \{COMP, RES\}$$

$$\{IN, COMP, RES\}$$

Figure 5.7: The Boolean lattice for the considered actor set

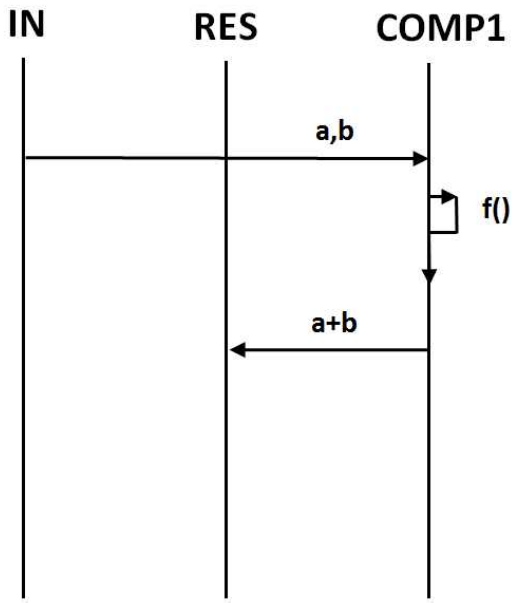

Figure 5.6: Our sample business process

Running our business process P, we obtain the following knowledge sets $K_S(t)$ for $t = 1, 2, 3$ (we omit the formal step $K_\emptyset(0) = K_\emptyset(1) = K_\emptyset(2) = \emptyset$):

The Process Initialization $t = 0$

$$\begin{array}{c} \{IN\}, \{COMP\}, \{RES\} \\ [INFO], [\phi], [\phi] \end{array}$$

$$\begin{array}{ccc} \{IN, COMP\} & & \{COMP, RES\} \\ [INFO] & \{IN, RES\} & [\phi] \\ & [INFO] & \end{array}$$

$$\begin{array}{c} \{IN, COMP, RES\} \\ [INFO] \end{array}$$

Figure 5.8: The knowledge sets at time $t = 0$

First Step $t = 1$

$$\begin{matrix}\{IN\},\{COMP\},\{RES\}\\ [INFO],[INFO],[\phi]\end{matrix}$$

$$\begin{matrix}\{IN,COMP\}\\ [INFO]\end{matrix} \qquad \begin{matrix}\{IN,RES\}\\ [INFO]\end{matrix} \qquad \begin{matrix}\{COMP,RES\}\\ [INFO]\end{matrix}$$

$$\begin{matrix}\{IN,COMP,RES\}\\ [INFO]\end{matrix}$$

Figure 5.9: The knowledge sets at time $t = 1$

Second Step $t = 2$

$$\begin{matrix}\{IN\},\{COMP\},\{RES\}\\ [INFO],[INFO],[INFO]\end{matrix}$$

$$\begin{matrix}\{IN,COMP\}\\ [INFO]\end{matrix} \qquad \begin{matrix}\{IN,RES\}\\ [INFO]\end{matrix} \qquad \begin{matrix}\{COMP,RES\}\\ [INFO]\end{matrix}$$

$$\begin{matrix}\{IN,COMP,RES\}\\ [INFO]\end{matrix}$$

Figure 5.10: The knowledge sets at time $t = 2$

The disclosure risk estimated by actor $IN1$ at $t = 0$ is zero, as there are no subsets $X \in 2^A$ such that $K_X(0) \neq \emptyset$ but the singleton $\{IN1\}$, whose only member coincides with the risk evaluating actor $IN1$.
At $t = 1$, however, there is another singleton such that $K_X(1) \neq \emptyset$ (in particular, $K_X(1) = INFO_1 \cup S_f$, namely the subset $X = \{COMP1\}$. All the other subsets for which $K_X(1) \neq \emptyset$ can be obtained by computing the ideal generated by {IN1,COMP1} w.r.t. the Boolean lattice's join ($\cup$), so their contribution to the risk estimation is zero (all their members had the same knowledge separately than they have when taken together).
The estimate by actor $IN1$ of disclosure risk in the part of $COMP1$ of information $INFO_1 \cup S_f$ at $t = 1$ can therefore be written as follows:

$$R(A_k, E_S) = R(IN1, E_{\{COMP1\}}) = P_{\{COMP1\}}(1) I_{\{COMP1\}}(IN1, 1) \tag{5.11}$$

where $P_{\{COMP1\}}(1)$ is the probability (assessed by $IN1$) that $COMP1$ will disclose at $t = 1$ the information it now holds, i.e. the data $INFO_1$ and the specification S_f of the local function $f()$ it computes (i.e. the addition). $I_{\{COMP1\}}(IN1, 1)$ is the resulting total damage to $IN1$ of the service provider $COMP1$ disclosing what it knows, i.e. the summands $INFO_1$ and the specification S_f.. At $t = 2$, another singleton such that $K_X(1) \neq \emptyset$ pops up, namely $X = \{RES\}$. Again, all the other subsets for which $K_X(1) \neq \emptyset$ can be obtained by computing the ideal generated by {IN1,COMP1,RES} w.r.t. the Boolean lattice's join ($\cup$) (in this case,

the entire lattice) so their contribution to the risk estimation is zero (all their members had the same knowledge separately than they have together).
Under the assumption the two disclosure events to be independent, total risk estimate at $t = 2$ by $IN1$ is therefore, as expected:

$$R(A_k, E_S) = R(IN1, E_{\{COMP1\}} \cup E_{\{RES\}}) \tag{5.12}$$

that becomes:

$$R(A_k, E_S) = P_{\{COMP1\}}(1) I_{\{COMP1\}}(IN1, 1) + P_{\{RES\}}(2) I_{\{RES\}}(IN1, 2) \tag{5.13}$$

Of course, risk as seen by other actors of P can also be evaluated by the same procedure: estimating the probability (or belief) that a disclosure event will occur as well as the damage they would incur should the disclosure event happen. For instance, risk estimated by RES at $t = 0$ is related to the singleton subset $\{IN1\}$ (the only one whose knowledge set is not empty; note that in this case, unlike before, it does not coincide with the risk assessor). We get:

$$R(A_k, E_S) = R(RES, E_{\{IN1\}}) = P_{\{IN1\}}(0) I_{\{IN1\}}(RES, 1) \tag{5.14}$$

In the same line, risk estimated by RES at $t = 1$ can be written as follows:

$$R(A_k, E_S) = R(RES, E_{\{IN1\}} \cup E_{\{COMP1\}}) \tag{5.15}$$

that becomes:

$$R(A_k, E_S) = P_{\{IN1\}}(0) I_{\{IN1\}}(RES, 1) + P_{\{COMP1\}}(1) I_{\{COMP1\}}(RES, 1) \tag{5.16}$$

Figure 5.11 shows the risk landscape associated to the protocol P for t=1.

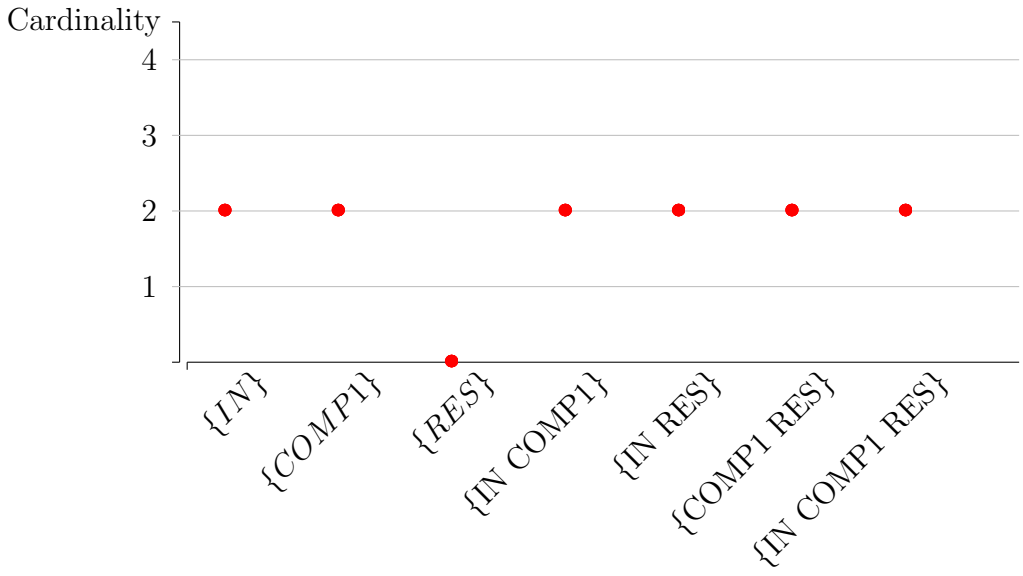

Figure 5.11: The risk landscape for the basic protocol

5.7.2 Decreasing disclosure risk

In order to alleviate risk, we need to apply our *risk management process* $M_{R(P)}$ where alternative strategies for dealing with the risks connected to threats are compared. The output of $M_{R(P)}$ will

be a *risk alleviation strategy*, which consists of modifications to P (including the deployment of security controls) that have the desired effect on the risk of executing it. While our methodology does not at present include specific guidance in the choice of such controls, we remark that the user can identify possible changes to P by searching pattern libraries offering alternative mechanisms for achieving and certifying the security properties of business process and services (see for instance [12]) [7].

We consider first a pattern of obfuscation of the local function $f()$. Instead of pushing the plaintext specification of S_f to the service provider $COMP1$, actor $IN1$ can use an obfuscation technique for computing the sum. While the obfuscation techniques themselves are outside the scope of this deliverable, we remark that a variety of obfuscation mechanisms have been proposed in the literature, including homomorphic encryption, evaluation of branching programs, and Garbled Circuits (GC). GC evaluation and homomorphic encryption are in principle both suitable for obfuscation of simple arithmetics operations like the one in our example. Let us assume that a GC technique is used for obfuscating addition (the size of the garbled adder circuit is small (linear in the size of the inputs), and its secure evaluation is efficient, as it is linear in the number of Oblivious Transfers (OT) and in the number of evaluations of a cryptographic hash function, for example SHA-256)[8]. In other words, in this representation of our business process P, our functional $G(f)$ denotes a Garbling Outsourcing Scheme (see Sect. 5.2.2) mechanism that locally (i.e., within its own trusted environment) computes a garbled function $G(f) = f'()$ corresponding to the sum and pushes the garbled specification of $f'()$, namely $S_{f'}$ to $COMP1$. The corresponding sample coreography is depicted in Figure 5.12.

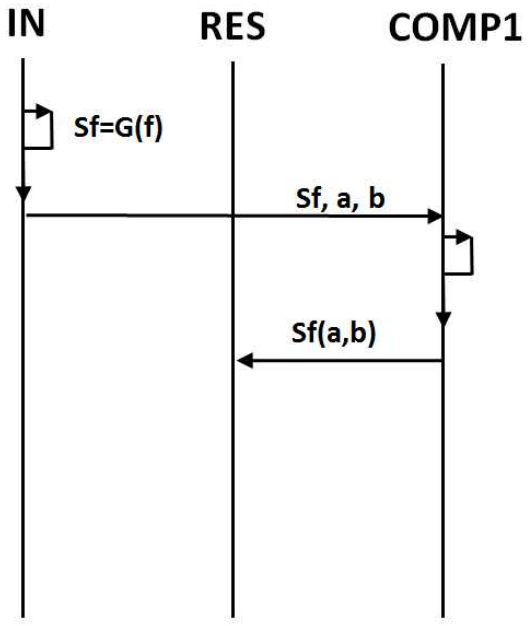

Figure 5.12: The coreography for the garbled sum computation

[7] Also, links between security properties and the corresponding threat spaces have been defined in the framework of several certification schemes [18].

[8] Garbled integer arithmetics has attracted much attention in the past few years [41] [36] both following Yao's original formulation and the alternative Goldreich-Micali-Wigderson (GMW) protocol. Also, [57] summarizes several depth-optimized circuit constructions for various standard arithmetic tasks.

The subset analysis carried out in Sect. 5.7.1 is now repeated after applying our modifications to the process P, but we can now apply the weaker version of our computation transparency assumption (Sect. 5.2). This way, the information disclosed to $COMP1$ does not include the local function specification any more. We get:

$$R(A_k, E_S) = R(IN1, E_{\{COMP1\}}) = P_{\{COMP1\}}(1) I'_{\{COMP1\}}(IN1) \tag{5.17}$$

where, since the knowledge reconstructible by $COMP1$ is now smaller than before, $I'_{\{COMP1\}} \leq I_{\{COMP1\}}$. The modification to P has therefore decreased risk; however, the amount of such decrease needs to be compared with the combined costs of (i) the local computation of garbling $G(f)$ on the part of $IN1$ and (ii) the additional complexity of computing the garbled function $f'()$ - instead of the original addition $f()$ - on the part of $COMP1$.

Another version of P that can be envisioned in order to decrease disclosure risk features *multiple service provisioning*, where the confidentiality of $INFO_1$ is increased by outsourcing the computation of $f()$ to multiple services, each getting to know only a portion (a share) of $INFO_1$. In this case of course we will need to extend our actor set to become $A = (IN1, COMP1, COMP2, RES)$ where, once more, at $t = 0$ actor $IN1$ holds the entire information item $INFO_1$ containing the two summands $INFO_1[1]$ and $INFO_1[2]$.
The power set 2^A of the actors is:

$$\{IN\}, \{COMP1\}, \{COMP2\}, \{RES\}$$

$$\{IN, COMP1\} \{IN, COMP2\} \{IN, RES\} \quad \{COMP1, COMP2\} \{COMP1, RES\} \quad \{COMP2, RES\}$$

$$\{IN, COMP1, COMP2\} \quad \{IN, COMP1, RES\} \quad \{IN, COMP2, RES\} \{COMP1, COMP2, RES\}$$

$$\{IN, COMP1, COMP2, RES\}$$

Figure 5.13: The Boolean lattice for the new actor set.

In words, the alternative version of our process P can be described as follows:

1. The input actor $IN1$ computes a local function to divide each summand into two shares, obtaining $INFO_1[1,1], INFO_1[1,2], INFO_1[2,1], INFO_1[2,2]$,.
2. $IN1$ sends $INFO_1[1,1]$ and $INFO_1[2,2]$ to $COMP1$, and $INFO_1[1,2]$ and $INFO_1[2,1]$ to $COMP2$
3. The two computation nodes compute a local function each on the shares they received, namely $f_{COMP1} = INFO_1[1,1] + INFO_1[2,2]$ and $f_{COMP2} = INFO_1[1,2] + INFO_1[2,1]$,
4. The two computation nodes send the results to the result node RES
5. RES computes $f_{RES} = f_{COMP1} + f_{COMP2}$ and outputs the result

For the sake of simplicity, let us assume for the moment that $IN1$ will generate two shares of $INFO_1$ using a naive technique, i.e. by taking respectively the Most Significant and the Least Significant Part (MSP-LSP) from the original value $INFO_1$.
For instance, if $INFO_1[1] = 25$ and $INFO_1[2] = 31$, then $COMP_1$ receives $INFO_1[1,1] = 20$ and $INFO_1[2,2] = 01$ and computes 21, while $COMP_2$ receives $INFO_1[1,2] = 05$ and $INFO_1[2,1] = 30$ and computes 35. Finally, RES receives 21 and 35 and computes 56
Of course, this simplified share generation would not really prevent $COMP$ nodes from guessing the original values, so our assumption of *Information completeness*: "Given a message delivery (A_s, A_d, m_{ts}), with $m_{ts} \leq INFO_s$, then $\beta_i(E_{sd}) = 0$" (see Sect. 5.2) should now be revised here to, say, $\beta_i(E_{sd}) = \frac{1}{10}$, assuming they both $COMP$ nodes know that the summands are two-figure integers. However, we will not deal with probability of autonomous guessing in this example, as the threat space we are considering involves collusions among multiple parties.
After defining the revised version of P and the underlying assumptions, we can estimate the knowledge sets corresponding to the new, secured version of the business process.

Initialization $t = 0$

$$\underset{INFO}{\{IN\}} \quad \underset{[\phi]}{\{COMP1\}} \quad \underset{[\phi]}{\{COMP2\}} \quad \underset{[\phi]}{\{RES\}}$$

$$\underset{[INFO]}{\{IN, COMP1\}} \underset{[INFO]}{\{IN, COMP2\}} \underset{[INFO]}{\{IN, RES\}} \underset{[\phi]}{\{COMP1, COMP2\}} \underset{[\phi]}{\{COMP1, RES\}} \underset{[\phi]}{\{COMP2, RES\}}$$

$$\underset{[INFO]}{\{IN, COMP1, COMP2\}} \underset{[INFO]}{\{IN, COMP1, RES\}} \underset{[INFO]}{\{IN, COMP2, RES\}} \underset{[\phi]}{\{COMP1, COMP2, RES\}}$$

$$\underset{[INFO]}{\{IN, COMP1, COMP2, RES\}}$$

Figure 5.14: The knowledge sets at time $t = 0$

First Step $t = 1$

$$\begin{array}{cccc} \{IN\} & \{COMP1\} & \{COMP2\} & \{RES\} \\ INFO & [INFO(1,1)] & [INFO(1,2)] & [\phi] \end{array}$$

$$\begin{array}{cccccc} \{IN,COMP1\} & \{IN,COMP2\} & \{IN,RES\} & \{COMP1,COMP2\} & \{COMP1,RES\} & \{COMP2,RES\} \\ [INFO(1,1)] & [INFO(1,2)] & [INFO] & [INFO(1,*)] & [INFO(1,1)] & [INFO(1,2)] \end{array}$$

$$\begin{array}{cccc} \{IN,COMP1,COMP2\} & \{IN,COMP1,RES\} & \{IN,COMP2,RES\} & \{COMP1,COMP2,RES\} \\ [INFO] & [INFO] & [INFO] & [INFO(1,*)] \end{array}$$

$$\begin{array}{c} \{IN,COMP1,COMP2,RES\} \\ [INFO] \end{array}$$

Figure 5.15: The knowledge sets at time $t = 1$

Second Step $t = 2$

$$\begin{array}{cccc} \{IN\} & \{COMP1\} & \{COMP2\} & \{RES\} \\ INFO & [INFO(1,1), INFO(2,2)] & [INFO(1,2), INFO(2,1)] & [\phi] \end{array}$$

$$\begin{array}{cccccc} \{IN,COMP1\} & \{IN,COMP2\} & \{IN,RES\} & \{COMP1,COMP2\} & \{COMP1,RES\} & \{COMP2,RES\} \\ [INFO] & [INFO] & [INFO] & [INFO] & [INFO(1,1), INFO(2,2)] & [INFO(2,1), INFO(1,2)] \end{array}$$

$$\begin{array}{cccc} \{IN,COMP1,COMP2\} & \{IN,COMP1,RES\} & \{IN,COMP2,RES\} & \{COMP1,COMP2,RES\} \\ [INFO] & [INFO] & [INFO] & [INFO] \end{array}$$

$$\begin{array}{c} \{IN,COMP1,COMP2,RES\} \\ [INFO] \end{array}$$

Figure 5.16: The knowledge sets at time $t = 2$

Third Step $t = 3$

$\{IN\}$ $\{COMP1\}$ $\{COMP2\}$ $\{RES\}$
$INFO$ $[INFO(1,1), INFO(2,2)]$ $[INFO(1,2), INFO(2,1)]$ $[\phi]$

$\{IN, COMP1\}$ $\{IN, COMP2\}$ $\{IN, RES\}$ $\{COMP1, COMP2\}$ $\{COMP1, RES\}$ $\{COMP2, RES\}$
$[INFO]$ $[INFO]$ $[INFO]$ $[INFO]$ $[INFO]$ $[INFO]$

$\{IN, COMP1, COMP2\}$ $\{IN, COMP1, RES\}$ $\{IN, COMP2, RES\}$ $\{COMP1, COMP2, RES\}$
$[INFO]$ $[INFO]$ $[INFO]$ $[INFO]$

$\{IN, COMP1, COMP2, RES\}$
$[INFO]$

Figure 5.17: The knowledge sets at time $t = 3$

Our risk estimate at $t = 1$ by $IN1$ is therefore:

$$R(A_k, E_S) = R(IN1, E_{\{COMP1\}} \cup E_{\{COMP1\}}) = P_{\{COMP1\}}(1) I_{\{COMP1\}}(IN1, 1) \tag{5.18}$$

where, again, $P_{\{COMP1\}}(1)$ is $IN1$'s estimated probability that $COMP1$ will disclose at $t = 1$ the information it now holds, and $I_{\{COMP1\}}(IN1, 1)$ is the resulting damage to $IN1$. We recall the assumption of *Information completeness* (Sect. 5.2): given a message delivery (A_s, A_d, m_{ts}), with $m_{ts} \leq INFO_s$, then $\beta_i(E_{sd}) = 0$. Once again this property expresses a zero estimated probability that the sharing generation scheme can be broken. Therefore $IN1$ will attribute no risk to this stage, where no subset of actors not including itself has knowledge of both $INFO_1$ shares. Once again, we remark that a weaker version of the *information completeness* assumption can be adopted here to reflect the evident weakness of the naive share generation scheme, which could be easily broken by $COMP1$ via an educated guess. However, the threat space under consideration does not include autonomous guesses, and the original assumption is kept. At $t = 2$, unlike the previous example, no other singleton exists such that $K_X(1) = INFO_1$. However, this time other subsets for which $K_X(1) = INFO_1$ can be obtained, namely $\{COMP1, COMP2\}$.
Under the assumption the two disclosure events to be independent, risk estimate at $t = 2$ by $IN1$ is therefore:

$$R(A_k, E_S) = R(IN1, E_{\{COMP1, COMP2\}}) \tag{5.19}$$

that becomes:

$$R(A_k, E_S) = P_{\{COMP1\}}(1) I_{\{COMP1\}}(IN1, 1) + P_{\{COMP1, COMP2\}}(2) I_{\{COMP1, COMP2\}}(IN1, 2) \tag{5.20}$$

$P_{\{COMP1, COMP2\}}$ is the probability that $\{COMP1, COMP2\}$ will actually share the information they have to reconstruct $INFO[1]$ times the damage $IN1$ would incur in, should the disclosure event actually happen. It is important to remark that, if this probability is considered null by default (for example, the assessor is sure that $COMP1$ and $COMP2$ do not know of each other, operate on different clouds or are not under the jurisdiction of the same regulatory authority) risk at $t = 2$ is also 0.

5.7.3 Comparing Alternative Processes via Risk Profiles

Our two alternative versions of process P have a different *risk profiles*, which we can now compare. In this Section we show how a quick version of this comparison can be carried out over visual representations of the two profiles, *even when we are not able to exactly quantify probabilities ad impacts*. In this case, we simply assign a conventional value to the impact of the disclosure of each knowledge set and represent the different risk profiles associated to different versions of P on the same plane, where the horizontal axis shows the subsets $S \in 2^A$ and the vertical one shows the impacts of the disclosure of K_S in the two versions of the process.
Of course, different versions of the same business process P will in general have different actor sets A. Figure 5.18 shows the Boolean lattices corresponding to our sample comparison; in this particular case, it is easy to see that there is a total embedding of the lattice corresponding to the original version of P into the lattice corresponding to the modified one:

{IN}, **{COMP1}**, $\{COMP2\}$, **{RES}**

{IN, COMP1} $\{IN, COMP2\}$ **{IN, RES}** $\{COMP1, COMP2\}$ **{COMP1, RES}** $\{COMP2, RES\}$

$\{IN, COMP1, COMP2\}$ **{IN, COMP1, RES}** $\{IN, COMP2, RES\}$ $\{COMP1, COMP2, RES\}$

$\{IN, COMP1, COMP2, RES\}$

Figure 5.18: The lattice embedding

Generally speaking, however, it is clear that there will be no complete embedding of one lattice into the other; rather, a *partial mapping* μ will have to be defined. In the visual representation of our profiles, subsets of the two lattices that are connected by μ will correspond the same entry in the horizontal axis. Figure 5.19 shows the risk profiles associated to the two competing versions of our sample business process P for t=1 to 4. The risk assessor can thus get a first visual representation of the difference between the profiles, to be later tuned with more accurate impact and probability assessments.

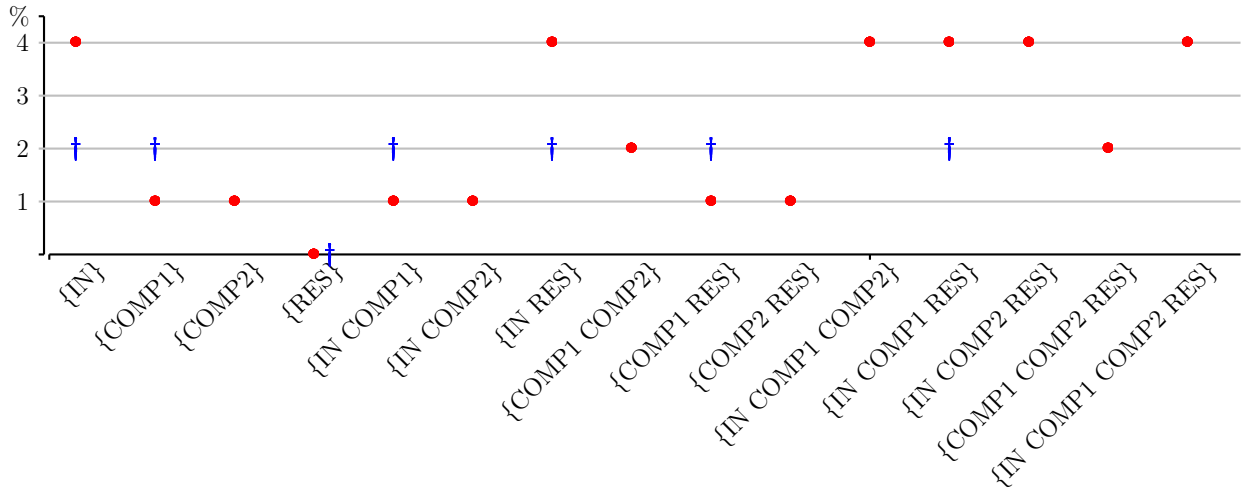

Figure 5.19: The risk profile of the secured P process compared with the original version

5.8 The Auction Scenario

Here we will introduce the rationale behind auctions, describe the main action mechanisms, point to the known security problems of auctions and then outline two auction scenarios exemplifying the computation of impacts and our approach.

5.8.0.1 The auction purpose, actors and process

Auction can provide efficient, distributed ways of solving goods and resource allocation problems [52, 60, 29, 66, 20, 40]. An auction consists of a set of potential *bidders* and an *auctioneer*, or bid taker (but there can be auctions with more than one bid taker). Auctions are usually discussed regarding situations where the auctioneer is the representative of a seller who wants to sell an item and get the highest possible payment for it while each bidder is a potential buyer or the representative of a potential buyer who wants to acquire the item at the lowest possible price. However, settings exist, where the auctioneer represents a buyer who wants to acquire a service at the lowest possible price and each bidder represents a seller who wants to offer the service at the highest possible payment. From the point of view of auction theory the two forms are totally analogous. Although auctions can and are used among cooperative agents, the key problem is how suitably designed auctions can work in systems consisting in self-interested agents. Auction theory analyzes protocols and agents' strategies in auctions. By looking for the strategies that self-interest agents will follow, pursues a main goal: designing interaction protocols (mechanisms) such that desirable social outcomes follow even though each agent acts based on self-interest (e.g. in terms of efficient allocation of goods/services). Auctions have acquired great relevance also over the internet thanks to commercial services such as e-bay, – which typically allows to auction tangible items and in cloud environments, where they can be used to auction the provision of services. Globally, they are called *e-auctions*.

The process An e-auction process involves several main activities (or tasks):

- Initialization: the auctioneer sets up the auction and advertises it (i.e., type of good, starting time, etc.).

- Registration: in order to participate in the auction, bidders must first register with the auctioneer (or a registration manager); this ensures that only valid bids are made and that bidders can be identified for payment purposes; registered bidders should be able to participate in any number of auctions rather than re-registering for each new auction.
- Bidding: a registered bidder computes his/her bid and submits it to the auctioneer (in some cases in a sealed envelope, in some other in an open declaration); the auctioneer checks the received bid to ensure that it conforms to the auction rules.
- Winner determination: the auctioneer determines the winner applying the auction rules (e.g. in an ordered set of bids, computing the first-best or the second-best).
- Winner communication: the auctioneer announces the winner to the parties.
- Contract issuing: the auctioneer issues to buyer and seller the contracts for the exchange of goods and payments. The contract obtained as the outcome of the auction is a legal document containing the information about the winner, the price the winning bidder has to pay and the item it has won if different items are being auctioned by the same seller. It has a seller's copy and a winner's copy. So, the seller can claim an amount of money equal to the winning price of an item by showing his copy of the contract to an appropriate authority, say a bank. On the other hand, the winner of a contract gets information about the item he won and the price he has to pay for it from the contract and can also legally claim the item by showing the contract to the appropriate person i.e. the seller. The enforcement of the contract is a key point in auctions: in case of weakly enforced contracts the auction can become object of several attacks aimed at gathering information and not at buying the auctioned item.

Economic requirements. The main economic requirement for an auction is *efficiency*. The auction performs two functions simultaneously: it determines the allocation of the item up for sale and it establishes the sale price. Accordingly it is natural to consider two aspects of auction performance: value maximization (efficiency) and revenue maximization: an auction is efficient if it allocates the item to the highest-value user and maximizes the total value in the process (being the total value the sum of the seller revenue and buyer profit); alternatively, if one takes purely the seller's point of view, revenue maximization is the primary goal. Factors that influence the efficiency of an auction scheme are costs, which include the computational and communication overhead required for registration, as well as costs for signing a bid, verifying a bid and winner determination. The two forms of efficiency requirements, in many situations, can be fulfilled at the same time. Auctions are typically evaluated using those two efficiency criteria, sometimes also considering auction implementation costs as a tie breaker.
E-auctions have great advantages over traditional auctions as auctioneer and bidders are not required to be physically present at the same location. This allows e-auctions to be much larger and more elaborate than traditional ones, i.e. potentially more efficient, However, it also provides more opportunities for the auction participants to cheat.

General security issues and requirements. Among the problems already present in traditional auctions that are still present in e-auctions we can mention the following (the problems can be more or less relevant depending on the different auction implementation):

- a buyer can cheat by colluding with other bidders to affect the settlement price, repudiating bids or failing to pay;

- the seller of the item could be in collusion with some of the buyers or might fail to deliver the goods;
- buyers/sellers could also forge a bid in an attempt to introduce fake bids in order to influence the auction operation;
- a corrupt auctioneer could award the auction to someone other than the legitimate winner;
- a bidder's personal information could be sold to marketing agencies, or used for malicious purposes.

Online auctions make some of the above concerns more relevant (in e-auctions, actors linked to the communication infrastructure become part to the game) and add further security concerns relating to bid privacy, bidder anonymity, correct evaluation and declaration of winner, etc. For instance, bidders may try to know the bid values of other bidders, based on weaknesses of the communication infrastructure; the auctioneer may not only try to know the bid values but also manipulate results (for example the knowledge of the second highest bid value helps to set reservation price in second price auction); incorrect outcome may result from introduction of false bids or modification of submitted bids, undue extension or shortening of bidding period and introduction of new bids based on information about submitted bids. For bidders, bid values may be sensitive information and loss of bid-privacy may reveal important information such as financial status etc. against their wishes.
It is well-known that current commercial e-auctions are inherently insecure. The corresponding sites only offer basic solutions that are designed to "clean up" after wrongdoing has taken place. All individuals are required to trust the auctioneer. However, cryptography can be used to solve some of these problems up-front. Security of various types of online auctions has received considerable attention from researchers during the past two decades [11, 9, 23, 43, 44, 47, 67, 8, 10, 33, 48, 61]. The main goals for a secure and anonymous electronic auction scheme are the following: *unforgeable bids* (if bids are forgeable a bidder can, for instance, be impersonated); *non-repudiation* (once a bidder has submitted a bid they must not be able to repudiate having made it: for example, if a bidder wins and does not want to pay, they might deny that they submitted the bid); *public verifiability* (there must be some publicly available information by which all parties can be verified as having correctly followed the auction protocol: this should include evidence of registration, bidding, and proof of winner/loser); *robustness* (for instance, the auction process must not be affected by invalid bids or by participants not correctly following the auction protocol).

5.8.0.2 Common auction mechanisms

Before introduction the four most common auction mechanisms – first-best and second-best sealed bid, open ascending and open descending bids – we remark that there are three qualitatively different auction settings, depending on how an agent's value of the item is formed: private value, common value and correlated value setting (for further details see for instance [37]). Given an auction mechanism the output of the auction depends on the settings. In private value auctions, the value of the good depends only on the agent's own preferences. An example is auctioning off public frequency bands in the television spectrum auction that the winning TV company will use to convey its own TV programs: the key is that the company will not resell the item, in which case the value would depend on the other agents' valuation. On the other hand in common value auctions an agent's value of an item depends entirely on the other agents' values of it, which are identical to the agent's by symmetry of this criterion.

For example auctioning treasury bonds fulfills this criterion: nobody inherently prefers having the bills and the value of the bill comes entirely from reselling possibilities. In correlated value auctions an agent value depends partly on its own preferences and partly on other's values. Hereafter we will always refer to the private value setting.

Open ascending-bid auctions (a.k.a.English auctions) The most well-known auction is probably the English auction, where the winner is the bidder with the highest bid and must pay an amount equal to this bid. Participants make increasingly higher bids, each stopping bidding when they are not prepared to pay more than the current highest bid. This continues until no participant is prepared to make a higher bid. The highest bidder wins the auction at the final amount bid. Eventually the auctioned good is only actually sold if the bidding reaches a reserve price set by the seller (the reserve price is the minimum bid that the principal is willing to recognize). An English auction is referred to as an open-bid auction because everyone knows the amount that a bidder has bidden.
In English auctions with a private value setting, every actor has an easy dominant strategy to play. For instance consider an auction with two bidders: if the auction is run with a price clock that raises the price continuously from zero until only one buyer is still willing to bid, then it is a weakly dominant strategy for the bidders to plan to keep bidding until the price reaches their valuation. Indeed, players who bid more than their valuations will take a loss if they win, while players who stop bidding before their valuations are reached are not taking advantage of a positive probability of making a gain. As a consequence of these optimal strategies the auctioned good will be sold to the bidder with the highest valuation. Notice, however, that the Seller will not be paid the winner valuation, because the auction is over when the loser stops bidding: the good is sold to the winner at the loser's valuation. With many bidders it would be sold to the winner at the valuation of the second-highest bidder.

Vickrey Auctions Second-price sealed-bid auctions (Vickrey auctions) as the other sealed auctions consists of two main stages: bidding and opening. During the bidding stage, bidders seal their bid (e.g., place it in an envelope) and submit it to the auctioneer. During the opening stage, the auctioneer opens all of the bids and determines the winner. The winner is the bidder with the highest bid. The winner is required to pay an amount equal to the second highest bid (i.e., the highest losing bid). In private value Vickrey auctions it is a weakly dominating strategy for both buyers to seal their true valuations of the auctioned good in to the envelopes they submit to the auctioneer. Whatever bids the other Buyers may have sealed in their envelopes, one can never benefit bidding below his true valuation because this can only lessen one's own probability of winning the auction without altering the amount paid in case of victory. Equally, one can never benefit from bidding above one's own true valuation, because this higher valuation will be useful to the victory only if another player has submitted a bid that is at least equal to one's own true valuation: in which case one would have to pay in case of victory at least the true valuation with the possibility to pay more.

English and Vickrey auctions are essentially the same in a private value, ideal setting: both elicit the truth from the buyers. In both cases, the optimal strategy is to bid v if the auctioned good valuation is v. It is straightforward to compute the expected payoffs for the participants.

Assume Sam is the seller of an item, Alice and Bob are the two buyers ($N = 2$) respectively with valuation $v = X_A$ and $w = X_B$ taken from a uniform distribution on a unit interval:

$v, w \in [0, 1]$. For sake of simplicity here assume that Sam's reserve price is $r = 0$. The probability that Alice wins (i.e. that her valuation is the maximum of the set of valuations) can be computed as follows: in the relevant interval, the cumulative of the distribution of the individual valuation is

$$F(v) = \int_0^v f(v')dv' = \int_0^v dv' = v$$

the cumulative of the maximum is $F_{\binom{max}{2}}(v) = F^2(v) = v^2$. The sought probability is

$$f_{\binom{max}{2}}(v) = 2v$$

.

The expected payment $m_A(v)$ for Alice in case of victory is given by v, times the expected value of Bob's valuation, i.e.

$$m_A(v) = \int_0^v f(w)w\,dw = \int_0^v w\,dw = \frac{1}{2}v^2$$

Alice expected gain is therefore $g_A(v) = v\,F(v) - m_A(v) = \frac{1}{2}v^2$. Since the problem is symmetric the same is true for Bob.

Because each buyer expects to pay $\frac{1}{2}v^2$ when they value the item at v, Sam's expected revenue from the sale is

$$R = 2\int_0^1 \frac{1}{2}v^2 dv = \frac{1}{3}$$

With N players the probability that a bidder with valuation $v = X_i$ wins is

$$f_{\binom{max}{N}}(v) = Nv^{(N-1)}$$

his expected payment $m_i(v)$ is given by the expected value of the the distribution of the maximum of $(N-1)$ players

$$m_i(v) = \int_0^v f_{\binom{max}{N-1}}(w)w\,dw = \frac{N-1}{N}v^N$$

his expected gain is finally

$$g_i(v) = v\,F(v) - m_i(v) = Nv^N - \frac{N-1}{N}v^N = \frac{N^2-N-1}{N}v^N$$

Sam's expected revenue from the sale is

$$R = N\int_0^1 \frac{N-1}{N}v^N\,dv = \frac{N-1}{N+1}$$

The above result holds for both English and Vickrey auction with private value. On the other hand, in correlated value auctions, the bids of other agents in the English auction provide information to the agent about its own valuation. Therefore English and Vickrey auctions are not strategically equivalent in general, and may lead to different results. Vickrey auctions have been widely advocated and adopted for use in computational multi-agent systems [50, 29, 66, 20, 40, 65, 51]. For example, versions of the Vickrey auctions have been used to allocate computation resources in operating systems [66, 20], to allocate bandwidth in computer networks [40, 65] and to computationally control building environments [29].

Dutch auctions In open descending-bid auctions (a.k.a. Dutch auctions), the price is set by the auctioneer at a level sufficiently high to deter all bidders, and is progressively lowered until a bidder, prepared to buy at the current price, stops the process, winning the auction. In a Dutch auction the bidder would be damaged from stopping the auction while the price remains above its valuation, but also stopping the auction when the price reaches its valuation will not bring any benefit to him. It is better for him to hold for a while hoping that no other buyers stops the process. The question of how much to stay below ones own valuation is the same that is posed in the case of first-price sealed bid auction: the reason is that the Dutch auction player might as well write down the price they plan to stop the auction before anything else happens. If they seal this prices in envelopes and hand them to the auctioneer running a first price sealed bid auction the result will be exactly the same as if they used them as stopping prices in a Dutch auction. It can be shown that in ideal symmetric buyer conditions the optimal strategy is to bid $\frac{1}{2}v$ if the auctioned good valuation is v. In Dutch auctions an agent bid matters only if it is the highest and no relevant information is revealed during the auction process.

First-price, sealed bid auctions First-price sealed-bid auctions consists of two stages: bidding and opening. During the bidding stage, bidders seal their bid and submit it to the auctioneer. During the opening stage, the auctioneer opens all of the bids and determines the winner. The winner is the bidder with the highest bid. The winner is required to pay an amount equal to its bid, i.e. the highest bid. In general there is no dominant strategy for acting in this auction. However with common knowledge assumptions regarding the probability distribution of the agents' values it is possible to determine Nash Equilibrium strategies for the agents (a Nash equilibrium for a game is the combination of individual strategies such that no agent has advantage in deviating from unilaterally).
For instance it can be shown that in ideal symmetric buyer conditions, both for the Dutch auction and the First-price sealed bid auction, the optimal strategy is to bid $\frac{1}{2}v$ if the auctioned good valuation is v. It can be shown however, that the expected payment is $\frac{1}{2}v^2$, the expected gain is $\frac{1}{2}v^2$.

So, English and Vickrey auctions do not shade the buyers' bids from the respective valuations at all. This happens, instead, with Dutch auctions and sealed bid first price auctions. However, all yield in ideal conditions the same output payoffs to the players. This outcome no longer holds as soon as the participants are for instance risk averse: In that case for instance he seller gains more in a Dutch auction than in an English auction. On the other hand, if the bidders valuations cease to be independent the seller can obtain more in an English auction than in a Dutch auction. Although the English auction mechanism is widely used in practice in Internet based auctions (eg. on E-Bay), Vickrey auctions have received wider attention from the computer science research community as they are less complicated to model compared to English auctions: the latter require a real time communication channel between auctioneer and bidders for price updates and winner determination, the former, on the other hand, only require a bidder to send a single message to the auctioneer. We will now focus on adaptations of the English and Vickrey auction scheme as online, cloud-based processes.

5.8.0.3 Impact from a collusion scenario in a simple auction

Hereafter we consider some issues that arise when a subset, or possibly all, of the bidders act collusively and engage in bid rigging with a view to obtaining lower prices. The resulting arrangement is called a *bidding ring* or *cartel*. While bidding rings are illegal, in real world

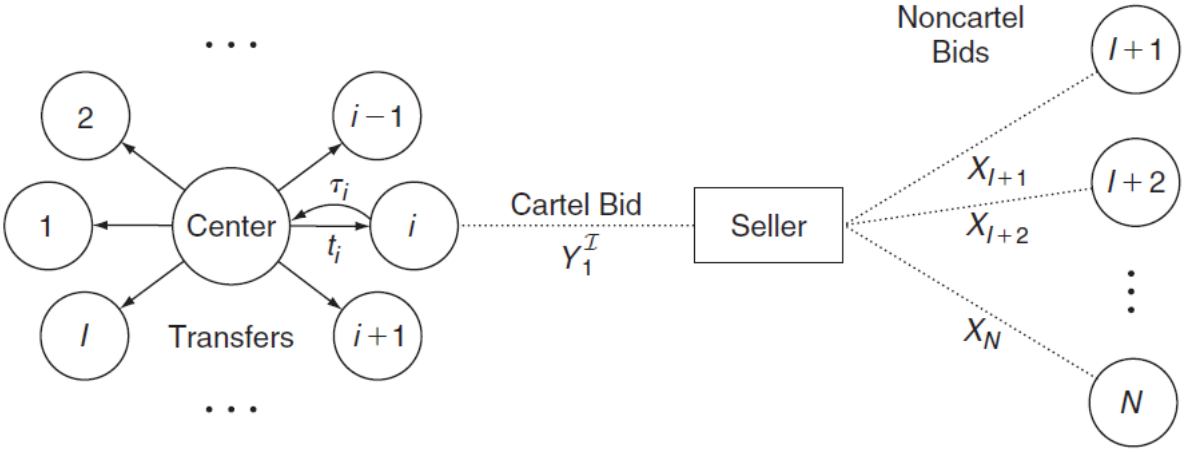

Figure 5.20: A bidding Ring

auctions they appear to be widely prevalent. Investigations of collusion in real world auctions constitute a significant component of antitrust activity [37]. In e-auctions the same issue is present if the bidders can know each other in real word, but can be present also when bidders who do not know each other get in contact exploiting the information leaked from the system. Theoretical models of collusion among bidders involve a mix of cooperative and noncooperative game theory: the former is needed to allocate to ring members the gain possibly obtained thanks to the the collusion: the problem can be faced for instance with an approach using the Shapley Values of the ring members. Here, we focus on the noncooperative part and take an example scenario from the on-line equivalent of a sealed-bid second-price auctions, which could be synthesized as follows.

Let us refer to the valuation of the bidders Alice, Bob and Carol by X_A, X_B, X_C respectively: in an honest second-price auction all the bidders bid their real valuation of the item. Suppose Alice and Carol have already submitted their valuations; Bob – who has not submitted his valuation yet – gets to know Alice's valuation and other contact information thanks to a leakage and discovers that $X_A > X_B$. Bob contacts Alice to establish a bidding ring: Bob will not bid X_B, but 0, so that in case of victory, Alice will not risk paying Bob's true valuation, but will pay Carol's true valuation: in case Carol's is lower than Bob's, this represents a gain $X_B - X_C$ for the ring, and a corresponding impact for the seller. We now extend this schema to N independent bidders, following the lines of [37].

Specifically, we suppose that each bidder's value is a random variable X_i which is distributed according to the cumulative distribution function F_i over some common interval $[0, 1]$. We remark that the assumption of a common interval is made only for notational convenience and can be easily relaxed. Let $\mathcal{I} \subseteq \mathcal{N}$ be the set of bidders in the bidding ring we indicate by $\mathcal{I} = \{1, 2, \ldots, I\}$ the set if their indexes and by $\mathcal{N} \setminus \mathcal{I} = \{I+1, I+2, \ldots, N\}$ the set of bidders outside the ring. For any set of bidders let the random variable Y_1^S denote the highest of the values of the bidders in S. It is in the interests of the bidding ring to make sure that the object goes in the hands of the ring member who values it the most, provided, of course, that it manages to win the object. In other words, in order to maximize profits the bidding ring must allocate the object efficiently among its members. But information regarding their values is privately held by the members and in order to function effectively, a bidding ring needs to gather this information and then to divide the gains from collusion among its members. How, and whether, both tasks can be accomplished is a key question.

In what follows we temporarily put aside the question of the internal functioning of a ring (for the second task, let us assume a suitable truth elicitation mechanism is used, e.g. in the form of another auction) and, assuming that it functions effectively, focus on the resulting gains and losses to the various parties. The presence of a ring in a second-price auction does not affect the behavior of bidders who are not members of the ring. It is still a weakly dominant strategy for a bidder $j \in \mathcal{I}$ to bid his or her value X_j. It is also weakly dominant for the ring to submit a bid equal to the highest value among its members, that is, $Y_1^{\mathcal{I}}$. Equivalently, we may think of the ring as being represented at the auction by the member with the highest value in the ring. The other members submit bids of 0, or if there is a reserve price, they bid at or below this price.

A bidding ring generates profits for its members, of course, by suppressing competition. Specifically, instead of N effective bids, only $N - I + 1$ effective bids are submitted since only one member of the cartel – the one with the highest value in the ring - submits a serious bid by bidding according to his or her value. The rest submit non-serious bids by bidding at or below the reserve price. The ring's profits come from the fact that, in certain circumstances, the price paid by a winning bidder from the ring is lower than it would be if there were no ring. Specifically, suppose that one of the ring members $i \in \mathcal{I}$ has a value X_i that is the highest of all bidders in the ring or outside the ring, i.e. $X_i = Y_1^{\mathcal{N}}$. Assuming that $X_i > r$, in the absence of a ring, this bidder would pay an amount equal to $P_i = max\{Y_1^{\mathcal{N}\setminus i}, r\}$ for the object. But if he were part of a functioning ring, his fellow members in $\mathcal{I}$ would bid at most r, so he would pay only

$$\widehat{P}_{\mathcal{I}} = max\{Y_1^{\mathcal{N}\setminus\mathcal{I}}, r\}$$

Thus, the expected payments of ring members are lower than they would be if the ring did not exist.
For a fixed reserve price r, let $m_i(X_i)$ denote the expected payment of bidder i with value x_i when there is no ring operating and all bidders behave non-cooperatively. Likewise, let $\widehat{m}_i(X_i)$ denote is expected payment when there is a ring, then

$$t_i(x_i) \equiv m_i(X_i) - \widehat{m}_i(X_i)$$

represents the contribution of bidder i to the rings expected profits when his value is x_i. The total ex ante expected profits of the ring amount to

$$t_{\mathcal{I}} \equiv \sum_{i\in\mathcal{I}} E[t_i(X_i)]$$

Notice that in the present context, the bidding ring exerts no externality whatsoever on bidders who are not part of the ring. First, the probability that a bidder who is not a member of the ring will win the object is the same whether or not the ring is functioning; in both cases it is just the probability that she has the highest value among all bidders. Furthermore, the price that a bidder $j \in \mathcal{I}$ would pay in the event that she wins is

$$max\{Y_1^{\mathcal{I}}, Y_1^{\mathcal{N}\setminus\mathcal{I}\setminus j}, r\} = max\{Y_1^{\mathcal{N}\setminus j}, r\}$$

the same as the price she would pay if there were no ring. Since for all bidders who are not part of the ring, neither the probability of winning nor the price upon winning is affected, the expected payments in the two situations are the same: the profits of these bidders are also unaffected. Since the profits of bidders outside the cartel are unaffected by its presence, the gains accruing to the cartel as a whole are equal to the loss suffered by the seller. This reasoning

also leads to the conclusion that the gains from collusion increase as the size of the ring increases.

The reasoning so far can be easily given a concrete quantitative exemplification in the case of uniform prior used previously. Assume as above that the common distribution of individual valuations is a uniform density $f(v)$ on $[0, 1]$ and the reserve price is $r = 0$. The expected payment of a player with valuation v in a non-rigged, second-price auction is given by the expected value of the the distribution of the maximum of $(N - 1)$ players (see above)

$$m_i(v) = \frac{N-1}{N} v^N$$

in a rigged second-price auction with a ring of I participants, instead, is given by the expected value of the the distribution of the maximum of $(N - I)$ players

$$\widehat{m}_{\mathcal{I}}(v) = \frac{N-I}{N-I+1} v^{N-I+1}$$

which is clearly lower. The expected advantage for the ring members translates in an expected damage for the seller

5.8.1 The Vickrey Auction Process Model

We are now ready to model a Vickrey sealed-bid auction using the approach of Sect.5.2. Bidders submit written bids without knowing the bid of the other people in the auction. The highest bidder wins, but the price paid is the second-highest bid. The Vickrey auction is close to eBay's system of proxy bidding; a slightly generalized version of it, named generalized second-price auction, is used in Google's and Yahoo!'s online advertisement programs [21, 64].
Figure shows our model of a Vickrey auction process VAP with two bidders. Node IN models the auctioneer, while a trusted $COMP$ node is used to compute via a suitable function $f()$ the second best bid. The RES node publishes both the amount to be paid (the second highest offer) and the winning bidder (the one who submitted the highest offer).
According to our methodology, the risk management process of VAP (M_{VAP}) first identifies the obfuscation mechanism for function $f()$. In principle, this risk alleviation could be relevant in this case even if the specification G_f of the Vickrey auction choice function is public from the start and accepted by all actors as part of the auction model. In principle, the obfuscation of the function, e.g. via GSC, can be tailored to hide from $COMP3$ whose offer is the second best, preventing it from learning more than the identity of the highest bidder and the amount of the second highest bid. However, the M_{VAP} comparison in this case largely coincides with the example in the previous section, and is therefore omitted.

A different alternative corresponds to a version of VAP where a Secure Multiparty Computation protocol is used to run the auction. The description below is the representation in our model of the approach described in [59]. After auctioneer IN has solicited their offers, the participating bidders $COMP1$ and $COMP2$ store their information items $INFO_i$ (their bids) in binary format , and the number of bits in each value is kept equal. in order to streamline the process representation avoiding (unrolling) loop, we have to choose a fixed value of this bit length (say 3) [9] Both bidders $COMP1$ and $COMP2$ send the most significant bit of their information items ($INFO_1$ and $INFO_2$) to $COMP3$. The latter actor computes a local function $f()$ (a

[9] According to the methodology, the analysis should in principle be repeated for all 64 possible pairs of bid values. It is however easy to see that in this process the risk profile is independent from any specific set of bids.

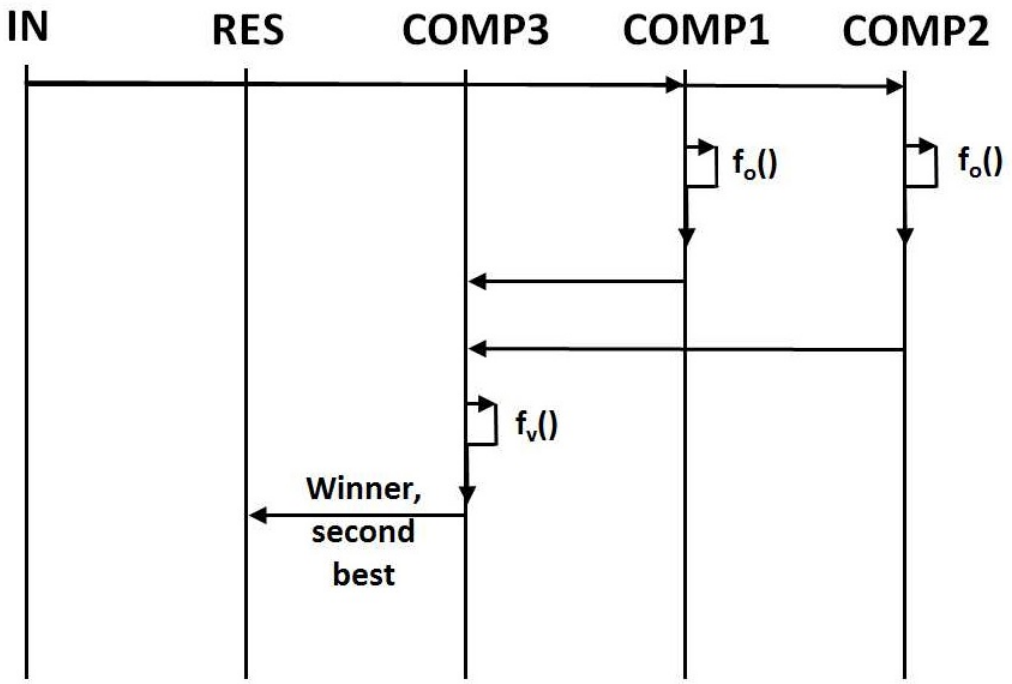

Figure 5.21: The Vickrey bid protocol

logical OR) of bits received. The result of $f()$ is sent from COMP3 to RES who publishes it. If the result of $f()$ is zero, $COMP3$ does nothing and waits for the next bit. If the result of $f()$ is a 1, all those parties who sent a 0 bit stop sending further bits to $COMP3$. It means those actors who sent a 1 bit continue sending the bits. When all the bits of the last party are sent, $COMP3$ publishes to RES (i) the winner (ii) the set of the results of the OR operations of the second-greatest value[10].

For the sake of conciseness, we do not show the entire Boolean lattice here. A quick visual comparison between the risk profiles of original and the modified version of the protocol VAP shows however that in the modified version the knowledge sets $K_{COMP1,COMP3}$ and $K_{COMP2,COMP3}$ do not increase the knowledge held by $COMP1$ and $COMP2$ until the tie-break moment, when the result is published anyway. This remains true even if we adopt a slightly weaker version of our *Information completeness assumption*, where knowing the MSB part of two information items allows to learn their relative order with respect to a domain total order relation.

Therefore, as intuition also suggests, the role of $COMP3$ can be safely assigned to an untrusted party (in terms of availability to "collude", i.e. to share with any of the bidders the bits of other bidders it receives) with no additional risk. It is important to remark that this result only holds for this specific version of VAP process with two bidders. Moving to a three bidders version with bidders $COMP1$, $COMP2$, $COMP3$ and compute node $COMP4$, the latter's availability at a given time $t = k$ to collude with, say, $COMP1$ by sharing the bit flows received from $COMP2$ and $COMP3$, together with the weaker version of our *Information completeness assumption* would allow $COMP1$ to change its $k+1$-th bit in order to keep up with competitors.

[10]In the case of this small example, the identity of the winner becomes known anyway. When multiple bidders are present and the privacy of the winner is to be preserved, RES can communicate the identity of the winner and the amount only to IN

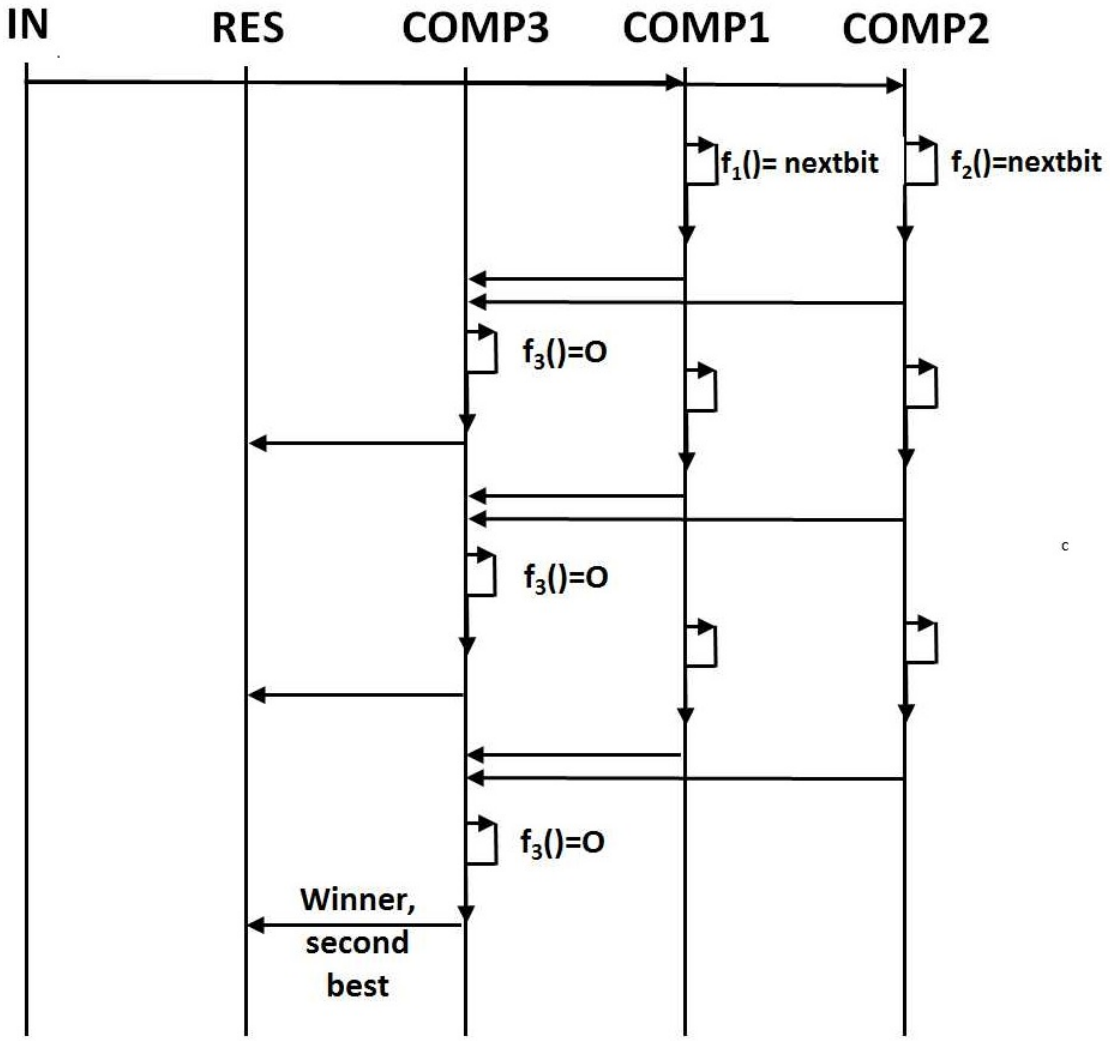

Figure 5.22: The alternative version of the Vickrey process

5.9 Conclusions

The risk analysis methodology presented in this deliverable covers some major activities of risk management involving the comparison of cloud-based process models including different security mechanisms from the point of view of the changes in risk they imply. While the methodology is still under evolution and refinement, especially as far as the scalability of the process models is concerned, we claim that our approach is extendable to cover many cost vs. risk assessment activities. Also, the methodology aims to gracefully extend existing machine-readable specification of processes like the W3C candidate recommendation for choreographies WS-CDL (http://www.w3.org/2002/ws/chor/) and to be supported by innovative software toolkits integrating existing choreography editors. Work on our risk assessment methodology, on our business process representation format and on its software support will continue during the next phases of the PRACTICE project

Bibliography

[1] Information risk analysis methodology IRAM. `https://www.securityforum.org/iram#iramtva`.

[2] ATOS. Risk analysis framework for a cloud specific environment, 2008.

[3] Information Systems Audit and Control Association. Cobit 5, 2013.

[4] RobertJ. Aumann and RogerB. Myerson. Endogenous formation of links between players and of coalitions: An application of the shapley value. In Bhaskar Dutta and MatthewO. Jackson, editors, *Networks and Groups*, Studies in Economic Design, pages 207–220. Springer Berlin Heidelberg, 2003.

[5] 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.

[6] Mihir Bellare, Viet Tung Hoang, and Phillip Rogaway. Foundations of garbled circuits. In Ting Yu, George Danezis, and Virgil D. Gligor, editors, *the ACM Conference on Computer and Communications Security, CCS'12, Raleigh, NC, USA, October 16-18, 2012*, pages 784–796. ACM, 2012.

[7] Dan Bogdanov, Liina Kamm, Sven Laur, and Pille Pruulmann-Vengerfeldt. Secure multiparty data analysis: end user validation and practical experiments. Cryptology ePrint Archive, Report 2013/826, 2013.

[8] Peter Bogetoft, Ivan Damgård, Thomas Jakobsen, Kurt Nielsen, Jakob Pagter, and Tomas Toft. A practical implementation of secure auctions based on multiparty integer computation. In *Financial Cryptography and Data Security*, pages 142–147. Springer, 2006.

[9] Colin Boyd and Wenbo Mao. *Security issues for electronic auctions*. Hewlett-Packard Laboratories, 2000.

[10] Phillip G Bradford, Sunju Park, Michael H Rothkopf, and Heejin Park. Protocol completion incentive problems in cryptographic vickrey auctions. *Electronic Commerce Research*, 8(1-2):57–77, 2008.

[11] Felix Brandt. Fully private auctions in a constant number of rounds. In *Financial Cryptography*, pages 223–238. Springer, 2003.

[12] Ingrid Buckley, Eduardo B. Fernández, Marco Anisetti, Claudio Agostino Ardagna, Seyed Masoud Sadjadi, and Ernesto Damiani. Towards pattern-based reliability certification of services. In *On the Move to Meaningful Internet Systems: OTM 2011 - Confederated International Conferences: CoopIS, DOA-SVI, and ODBASE 2011, Hersonissos, Crete, Greece, October 17-21, 2011, Proceedings, Part II*, pages 560–576, 2011.

[13] Daniele Catteddu and Giles Hogben. Cloud computing: Benefits, risks and recommendations for information security. Technical report, ENISA, 2009.

[14] Ann Cavoukian. Privacy risk management: Building privacy protection into a risk management framework to ensure that privacy risks are managed by default. Technical report, Information and Privacy Commissioner - Ontario - Canada, 2010.

[15] T. Chen. *Information and Risk Management*. 2009.

[16] CISCO. Data leakage worldwide white paper: The high cost of insider threats, 2011.

[17] Cloud Security Alliance. Security guidance for critical areas of focus in cloud computing v2.1, 2009.

[18] Ernesto Damiani, Claudio Agostino Ardagna, and Nabil El Ioini. *Open Source Systems Security Certification*. Springer, 2009.

[19] Folker den Braber, Gyrd Brndeland, Heidi E. I. Dahl, Iselin Engan, Ida Hogganvik, Mass S. Lundand Bjrnar Solhaug, Ketil Stlen, and Fredrik Vraalsen. The coras model-based method for security risk analysis. Technical report, SINTEF, 2006.

[20] K Eric Drexler and Mark S Miller. Incentive engineering for computational resource management. *The ecology of Computation*, 2:231–266, 1988.

[21] Benjamin Edelman and Michael Schwarz. Internet advertising and optimal auction design. In Ying Li, Bing Liu, and Sunita Sarawagi, editors, *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Las Vegas, Nevada, USA, August 24-27, 2008*, page 1. ACM, 2008.

[22] Josep Oriol Fitó and Jordi Guitart. Introducing risk management into cloud computing. Technical Report UPC-DAC-RR-2010-33, Technical University of Catalonia, 2010.

[23] Matthew K Franklin and Michael K Reiter. The design and implementation of a secure auction service. *Software Engineering, IEEE Transactions on*, 22(5):302–312, 1996.

[24] Sailesh Gadia. Cloud computing risk assessment: A case study. *ISACA Journal*, (1):1–6, 2012.

[25] The Open Group. Risk taxonomy, 2008.

[26] Jay Heiser and Mark Nicolett. Assessing the security risks of cloud computing, 2008.

[27] T. Hoomans, J. Seidenfeld, A. Basu, and D. Meltzer. Systematizing the use of value of information analysis in prioritizing systematic reviews. Technical Report 12-EHC109-EF, Agency for Healthcare Research and Quality, 2012.

[28] Ronald A. Howard. Information value theory. *IEEE Trans. Systems Science and Cybernetics*, 2(1):22–26, 1966.

[29] Bernardo A Huberman and Scott H Clearwater. A multi-agent system for controlling building environments. In *ICMAS*, pages 171–176, 1995.

[30] ISO. ISO/IEC 27002:2005 information technology - security techniques - code of practice for information security management, 2005.

[31] ISO. ISO31000:2009, risk management - principles and guidelines. 2009.

[32] ISO. ISO31010:2009, risk management - risk assessment techniques. 2009.

[33] Ari Juels and Michael Szydlo. A two-server, sealed-bid auction protocol. In *Financial Cryptography*, pages 72–86. Springer, 2003.

[34] Burton S. Kaliski, Jr. and Wayne Pauley. Toward risk assessment as a service in cloud environments. In *Proceedings of the 2Nd USENIX Conference on Hot Topics in Cloud Computing*, HotCloud'10, pages 13–13, Berkeley, CA, USA, 2010. USENIX Association.

[35] A.U. Khan, M. Oriol, M. Kiran, Ming Jiang, and K. Djemame. Security risks and their management in cloud computing. In *Cloud Computing Technology and Science (CloudCom), 2012 IEEE 4th International Conference on*, pages 121–128, Dec 2012.

[36] Vladimir Kolesnikov. Gate evaluation secret sharing and secure one-round two-party computation. In Bimal K. Roy, editor, *Advances in Cryptology - ASIACRYPT 2005, 11th International Conference on the Theory and Application of Cryptology and Information Security, Chennai, India, December 4-8, 2005, Proceedings*, volume 3788 of *Lecture Notes in Computer Science*, pages 136–155. Springer, 2005.

[37] Vijay Krishna. *Auction theory*. Academic press, 2009.

[38] Antonio Kung, Alberto Crespo Garcia, Nicols Notario McDonnell, Inga Kroener, Daniel Le Mtayer, Carmela Troncoso, Jos Mara del lamo, and Yod Samuel Martns. Pripare: A new vision on engineering privacy and security by design. Technical report, PRIPARE, 2014.

[39] Robert La Rose and Nora J. Rifon. Promoting i-safety: Effects of privacy warnings and privacy seals on risk assessment and online privacy behavior. *Journal of Consumer Affairs*, 41(1):127–149, 2007.

[40] Jeffrey K MacKie-Mason and Hal R Varian. Pricing the internet. Technical report, Econ-WPA, 1994.

[41] Dahlia Malkhi, Noam Nisan, Benny Pinkas, and Yaron Sella. Fairplay - secure two-party computation system. In Matt Blaze, editor, *Proceedings of the 13th USENIX Security Symposium, August 9-13, 2004, San Diego, CA, USA*, pages 287–302. USENIX, 2004.

[42] Ronald K Mitchell, Bradley R Agle, and Donna J Wood. Toward a theory of stakeholder identification and salience: Defining the principle of who and what really counts. *Academy of management review*, 22(4):853–886, 1997.

[43] Moni Naor, Benny Pinkas, and Reuban Sumner. Privacy preserving auctions and mechanism design. In *Proceedings of the 1st ACM conference on Electronic commerce*, pages 129–139. ACM, 1999.

[44] Khanh Quoc Nguyen and Jacques Traoré. An online public auction protocol protecting bidder privacy. In *Information Security and Privacy*, pages 427–442. Springer, 2000.

[45] NIST. Federal information processing standard (fips) 65, guideline for automatic data processing risk analysis, 1979.

[46] NIST. Recommended security controls for federal information systems and organizations, 2009.

[47] Kazumasa Omote and Atsuko Miyaji. A practical english auction with one-time registration. In *Information Security and Privacy*, pages 221–234. Springer, 2001.

[48] David C Parkes, Michael O Rabin, and Christopher Thorpe. Cryptographic combinatorial clock-proxy auctions. In *Financial Cryptography and Data Security*, pages 305–324. Springer, 2009.

[49] Michael O. Rabin. How to exchange secrets with oblivious transfer. *IACR Cryptology ePrint Archive*, 2005:187, 2005.

[50] Kate Reynolds. Going, going, gone! a survey of auction types, 1996.

[51] Jeffrey S Rosenschein. *Rules of encounter: designing conventions for automated negotiation among computers*. MIT press, 1994.

[52] Tuomas Sandholm. An implementation of the contract net protocol based on marginal cost calculations. In *AAAI*, volume 93, pages 256–262, 1993.

[53] Amit Sangroya, Saurabh Kumar, Jaideep Dhok, and Vasudeva Varma. Towards analyzing data security risks in cloud computing environments. In Sushil K. Prasad, Harrick M. Vin, Sartaj Sahni, Mahadeo Jaiswal, and Bundit Thipakorn, editors, *Information Systems, Technology and Management - 4th International Conference, ICISTM 2010, Bangkok, Thailand, March 11-13, 2010. Proceedings*, volume 54 of *Communications in Computer and Information Science*, pages 255–265. Springer, 2010.

[54] P. Saripalli and B. Walters. Quirc: A quantitative impact and risk assessment framework for cloud security. In *Cloud Computing (CLOUD), 2010 IEEE 3rd International Conference on*, pages 280–288, July 2010.

[55] Grant T Savage, Timothy W Nix, Carlton J Whitehead, and John D Blair. Strategies for assessing and managing organizational stakeholders. *The executive*, 5(2):61–75, 1991.

[56] 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.

[57] Thomas Schneider and Michael Zohner. GMW vs. yao? efficient secure two-party computation with low depth circuits. In Ahmad-Reza Sadeghi, editor, *Financial Cryptography and Data Security - 17th International Conference, FC 2013, Okinawa, Japan, April 1-5, 2013, Revised Selected Papers*, volume 7859 of *Lecture Notes in Computer Science*, pages 275–292. Springer, 2013.

[58] A.S. Sendi and M. Cheriet. Cloud computing: A risk assessment model. In *Cloud Engineering (IC2E), 2014 IEEE International Conference on*, pages 147–152, March 2014.

[59] R. Sheikh and D.K. Mishra. Protocols for getting maximum value for multi-party computations. In *Mathematical/Analytical Modelling and Computer Simulation (AMS), 2010 Fourth Asia International Conference on*, pages 597–600, May 2010.

[60] R Smith. Communication and control in problem solver. *IEEE Transactions on computers*, 29:12, 1980.

[61] Koutarou Suzuki and Makoto Yokoo. Secure generalized vickrey auction using homomorphic encryption. In *Financial Cryptography*, pages 239–249. Springer, 2003.

[62] The Economist Intelligence Unit. Managing business risks in the information age, 1998.

[63] Paolo Trucco, Enrico Cagno, Fabrizio Ruggeri, and Ottavio Grande. A bayesian belief network modelling of organisational factors in risk analysis: A case study in maritime transportation. *Rel. Eng. & Sys. Safety*, 93(6):845–856, 2008.

[64] Hal R. Varian. Position auctions. *International Journal of Industrial Organization*, 25(6):1163 – 1178, 2007.

[65] William Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *The Journal of finance*, 16(1):8–37, 1961.

[66] Carl A Waldspurger, Tad Hogg, Bernardo A. Huberman, Jeffrey O. Kephart, and W. Scott Stornetta. Spawn: A distributed computational economy. *Software Engineering, IEEE Transactions on*, 18(2):103–117, 1992.

[67] Changjie Wang and Ho-fung Leung. Anonymity and security in continuous double auctions for internet retails market. In *System Sciences, 2004. Proceedings of the 37th Annual Hawaii International Conference on*, pages 10–pp. IEEE, 2004.

[68] Vic Winkler. Cloud computing: Risk assessment for the cloud. *Technet Magazine*, January 2012.

[69] David Wright. Should privacy impact assessments be mandatory? *Commun. ACM*, 54(8):121–131, 2011.