www.jusletter.ch

Alexandre Jotterand

Personal Data or Anonymous Data: where to draw the lines (and why)?

The concept of personal data is as complex as it is important. This article analyzes key legal issues surrounding the concept. It exposes the flaws of the «absolute vs relative» approaches regularly used by legal scholars to address the issue of identifiability and proposes a new test, duly taking into account the environment in which the data is processed. The article further analyzes the handling and transfer of pseudonymized data in light of the decisions of the Federal Tribunal and of cantonal courts, and the resulting obligations for both the data provider and data recipient.

Category of articles: Articles
Field of Law: Data Protection

Citation: Alexandre Jotterand, Personal Data or Anonymous Data: where to draw the lines (and why)?, in: Jusletter 15 August 2022

ISSN 1424-7410, jusletter.weblaw.ch, Weblaw AG, info@weblaw.ch, T +41 31 380 57 77

Contents

1. Legal Framework

[1] The processing of personal data in Switzerland by private and federal public bodies is primarily governed by the Swiss Federal Act on Data Protection (FADP)[1] and its ordinance (DPO).[2] The revised FADP[3] (**revFADP**) is expected to enter into force in September 2023.[4] Due to Swiss federalism, the processing of personal data by public bodies is not governed by the FADP, but by the data protection laws of each canton. The EU General Data Protection Regulation (GDPR)[5] may also apply to processing activities performed in Switzerland pursuant to Art. 3 GDPR.[6]

[2] In addition to the legislation cited above, various sector-specific federal laws and regulations set forth requirements regarding the use of personal data in specific circumstances. As a notable example, which will be addressed in more detail in this analysis, research on human beings, and more specifically research on human diseases and on the structure and functioning of the human body,[7] is governed in Switzerland primarily by the Federal Act on Research involving Human Beings (HRA)[8] and its ordinances, including the Ordinance on Human Research with the Exception

1 Federal Act on Data Protection (FADP) of 19 June 1992, RS 235.

2 Ordinance to the Federal Act on Data Protection (DPO) of 14 June 1993, RS 235.11.

3 Loi fédérale sur la protection des données (LPD) du 25 septembre 2020, FF 2020 7373.

4 The final version of the revised federal data protection ordinance is not known at the time of writing.

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

6 Regarding the application of the GDPR, see Alexandre Jotterand/Frédéric Erard, Recherche sur l'être humain et données personnelles, in: Jusletter 30 August 2021, §§ 12–17.

7 Art. 2 para. 1 HRO.

8 RS 810.30.

of Clinical Trials (HRO).[9] The HRA is a special law (*lex specialis*) defining special rules for data processing in the context of research on human beings.[10] Its provisions thus take precedence over the «general» data protection laws (FADP and cantonal data protection laws), without, however, completely replacing them. General data protection laws will apply where the HRA is silent, for example in the application of general data protection principles (e.g. the principles of proportionality or legality). The processing of personal data in the context of research projects will therefore be subject to the provisions of the HRA, in addition to the general regulations of data protection law. In particular, the HRA contains specific provisions on the right to information and the obligation to obtain the consent of the research participants (in particular, Articles 7, 8, and 16 to 18 HRA), as well as a specific chapter on the conditions for the re-use of personal data related to health (Chapter 4, Articles 32–35 HRA).

2. The Concept of «Personal Data»

[3] The concept of personal data – **any information relating to an identified or identifiable person**[11] – has been a cornerstone of data protection laws since their very inception, both in Switzerland and the EU.[12] Data protection laws follow a dichotomous model: data is either personal data or anonymous information, and the laws will only apply in case there is a *processing* of *personal data*. If the data does not qualify as personal data – because it is technical by nature, or because it cannot be linked to an individual – data protection laws will not apply.

[4] It is generally accepted that the concept of personal data must be understood broadly and should also include data with a very low personal reference and a low risk to the personality of the data subject.[13] However, despite the crucial importance of the notion (as it determines the rules that apply to the data), its boundaries remain blurry and debated.

[5] There are four components to the concept of personal data[14]:

(a) an information;
(b) a (natural) person;

---

9 RS 810.301. The other ordinances to the HRA are the Ordinance on Clinical Trials with the exception of Clinical Trials of Medical Devices (**ClinO**; RS 810.305) and the Ordinance on Clinical Trials with Medical Devices (**ClinO-MD**; RS 810.306).

10 SHK HFG-Brunner, Vorbemerkungen Art. 56–61 N 4, in: Bernhard Rütsche (ed.), Humanforschungsgesetz (HFG), Berne 2015; Frédéric Erard, Les données codées dans le contexte de la recherche : personnelles ou anonymes, AJP/PJA 2021/5, p. 613; David Rosenthal, Die rechtlichen und gefühlten Grenzen der Zweitnutzung von Personendaten, sic! 2021, p. 168, p. 170.

11 Art. 3(a) FADP. This concept is defined in similar terms in the revFADP (Art. 5(a) revFADP) and the GDPR (Art. 4(1) RGPD), as well as in both the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data of 28 January 1981 (Convention 108), Art. 2(a), and its revised version of 18 May 2018 (Convention 108+). For the predecessor to the GDPR, see also Art. 2(a) of the Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data (DPD). 24 October 1995, OJ L 281/31.

12 For the EU, see Lorenzo Dalla Corte, Scoping personal data: Towards a nuanced interpretaton of the material scope of EU data protection law, European Journal of Law and Technology, Vol 10, Issue 1, 2019.

13 Judgement of the Commercial Court (Handelsgericht) of the Canton of Zurich HG190107O of 4 May 2021, para. 3.2.3 a); Judgement of the High Court (Obergericht) of the Canton of Zurich PP190037-O/U of 30 January 2020, para. 4.3.2; Michèle Finck/Frank Pallas Distinguishing personal from non-personal data, in: International Data Privacy Law, Volume 10, Issue 1, February 2020, pp. 11–36, p. 13.

14 Decision of the Appellationsgericht des Kantons Basel-Stadt ZB.2019.3 of 9 August 2019, para. 4.2.1. Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007, p. 6.

(c) a link («relating to») connecting the information to the data subject; and
(d) identifiability.

[6] The boundaries of the concept of personal data depend on the interpretation of, and the relation between, those four components.

[7] We will focus our analysis on the final component, «*identifiability*». The first two components («*information*» and «*person*») generally pose little difficulty, while the third one (that the information must «*relate to*» a «person») can be complex. Analyzing it in detail would require a separate article. We will, however, present its key characteristics.

[8] The concept of «*information*» is not defined by data protection laws (nor is «data» which is treated as a synonym).[15] The term refers to any type of information regardless of its content and form (e.g. a sign, word, image, sound or a combination thereof), storage method, or the designation of the data file.[16] This includes information «stored» in the memory of a person.[17]

[9] The concept of «*person*» is defined by each law. Under the current FADP (and current cantonal data protection laws), this includes both individuals and legal entities. The GDPR and revFADP only include individuals.[18]

[10] The information «*relates to*» a person when there is a relational link with that person. This is sometimes easy to assess, for instance in the case of the results of a patient's medical test contained in her medical records). In such case, the information is directly *about* that person. In practice, however, there are many situations where it is not easy to determine if the information indeed relates to an individual in the sense given above. The link between the information and the individual may be tenuous and there may be many ways to satisfy the requirement for a relational link between data and the person.[19] Information about objects, events and processes may also be related to a person who has a certain relationship to the object, event or process, or has a certain influence on them (e.g. the price of a real estate property conveys information about its owner). Thus, even if information is not directly about an individual (***content***), it will still be considered personal data if it can reasonably be expected to be used to evaluate or influence an individual (***purpose***) or if its use may have an impact on the individual (***outcome***).[20] The extent to which the information relates to the personality of the data subject is irrelevant.[21]

3. The Concept of «Identifiability»

3.1. In General

[11] To qualify as personal data, the data must relate to an **identified** or **identifiable** person.

15 Dalla Corte (n 12), p. 3.

16 ATF 147 III 139, para. 3.1; ATF 136 II 508 = JdT 2011 II 446, para. 3.2; Zurich, HG190107O (n 13, para. 3.2.3 a).

17 ATF 147 III 139, para. 3.4. The access right will, however, not apply to personal data possibly stored in the brain among the «ordinary memories of a person» («*gewöhnlichen Erinnerungen*»).

18 We will focus in this article on the concept of personal data relating to individuals, to the exclusions of legal entities.

19 Dalla Corte (n 12), p. 1.

20 Groupe Article 29 Opinion 4/2007 (n 14), pp. 10–11.

21 Basel-Stadt ZB.2019.3 (n 14), para. 4.2.1 and the references cited. See also ATF 138 II 346 para. 6.1.

[12] Identification results from the interpretation of ***identifiers*** present in the information or that can be deduced from such information or its context (including any *information about the information*, i.e. metadata). *Identifiers* may include the name, an address, an identification number, location data, an online identifier or factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of a person.[22] Identifiers may directly refer to the person, e.g. her age, profession, nationality, or indirectly refer to her, e.g. in case of information that relates to the family members of the data subject. The notion of identifier is not limited: anything that may allow, directly or indirectly, to identify a person is an identifier.

[13] A person is **identified** if it is clear from the information itself that this is precisely the person concerned (e.g. ID document).[23]

[14] The person is **identifiable** when she cannot be clearly identified from the data alone, but can be identified from the circumstances, i.e. **from the context of the information or on the basis of additional information** (for example, where the identity of an owner can be determined from information about her real estate property):

> *A person is identifiable if she is not clearly identified by the data alone, but her identity can be inferred from the circumstances, i.e. from the context of a piece of information or on the basis of additional information (e.g. when a property owner can be identified from information about real estate).* [24]

[15] It is not necessary for the entire dataset to be identifiable. Identifiability of part of the information is sufficient:

> *Furthermore, identifiability is to be affirmed if it relates to at least part of the stored information.*[25]

[16] The threshold for identifiability is low. The decisive factor is that the information can be attributed to one or more persons.[26] The boundaries of identifiability are, however, heavily debated, as we will see below.

---

22 Federal Council Dispatch Concerning the Federal Law on the Total Revision of the Federal Law on Data Protection And on the Amendment of Other Federal Laws, FF 2017 6565 p. 76; Art. 4 para. 1 GDPR.

23 ATF 138 II 346 para. 6.1 = JdT 2013 I p. 71, 77; Judgment of the Swiss Federal Tribunal 4A_365/2017 of 26 February 2018, para. 5; Philippe Meier, Protection des données. Fondements, principes généraux et droit privé, Bern 2011, N 431 p. 201.

24 ATF 138 II 346 (unofficial translation of: «Bestimmbar ist die Person, wenn sie zwar allein durch die Daten nicht eindeutig identifiziert wird, aus den Umständen, das heisst aus dem Kontext einer Information oder aufgrund zusätzlicher Informationen auf sie geschlossen werden kann (z.B. wenn aus Angaben über Liegenschaften der Eigentümer ausfindig gemacht werden kann»).

25 ATF 138 II 346 para. 6.1 (unofficial translation of: «Weiter ist die Bestimmbarkeit zu bejahen, wenn sie sich zumindest auf einen Teil der gespeicherten Informationen bezieht»).

26 ATF 136 II 508, para. 3.2 (unofficial translation of «*Entscheidend ist, dass sich die Angaben einer oder mehreren Personen zuordnen lassen*»).

3.2. Moving Away From The «Absolute Vs. Relative» Approaches of Identifiability

[17] There is a dispute among scholars – both in Switzerland[27] and in the EU[28] – on the approach of identifiability that must be followed. This is often framed as the «absolute vs. relative approach» dispute, according to which there are two conflicting approaches to assess the identifiability of information (and thus of classifying personal and non-personal [or anonymized] data):[29]

> (a) Under the **absolute approach**, information will be considered as personal data **if there is any theoretical possibility of identification by anyone**. This includes two different components:
>
> - **component 1** (means): all ways and means to identify must be considered, without any regard to expense, etc. (a theoretical chance of re-identification is sufficient); and
> - **component 2** (actors): it is sufficient if anyone in the world can identify the individual: if the information is personal data for someone, it will be considered as personal data for everyone.
>
> (b) Conversely, under the **relative approach**, data is personal data only if there are realistic chances of identification (= component 1) by the data holder (= component 2).

[18] Swiss scholars seem to predominantly advocate the relative approach.[30] However, the debate hinges on concepts that are imprecise and too schematic. As stated above, each approach is composed of two separate components, and the focus of the authors taking position in this debate is not always clear (i.e. a *strict approach* considering the two components cumulatively, or a *more nuanced approach* focusing only on one of them).[31] Furthermore, different stances or variations may be followed for each component, resulting in additional forms of *nuanced approaches*. This is

---

27 Célian Hirsch/Emilie Jacot-Guillarmod, Les données bancaires pseudonymisées – Du secret bancaire à la protection des données, in: RSDA 2020 p. 151, pp. 160–161.

28 James Scheibner et al., Revolutionizing Medical Data Sharing Using Advanced Privacy-Enhancing Technologies: Technical, Legal, and Ethical Synthesis, J Med Internet Res. 2021 Feb 25, p. 6; Gerald Spindler/Philipp Schmechel, Personal data and encryption in the European General Data Protection Regulation, J Intellect Prop Inf Technol Electron Commer Law 2016; 7(2):163.

29 Meier (n 23), p. 445.

30 Andrea Martani/Philipp Egli/Michael Widmer/Bernice Elger, Data protection and biomedical research in Switzerland: setting the record straight, in: Swiss Medical Weekly 2020;150:w20332, p. 3; Sylvain Métille, L'utilisation de l'informatique en nuage par l'administration publique, AJP/PJA 6/2019 p. 609, p. 615; David Rosenthal, Das neue Datenschutzgesetz, in: Jusletter 16 November 2020; David Rosenthal, Personendaten ohne Identifizierbarkeit, in: digma 2017/4, pp. 198 ff, p. 202; David Rosenthal/Yvonne Jöhri, Handkommentar zum Datenschutzgesetz, Zurich 2008 (quoted. Author, Handkommentar DSG), Art. 3 let. a N 36; Beat Rudin, in: Bruno Baeriswyl/Kurt Pärli (éd.), Datenschutzgesetz (DSG), Berne 2015, Art. 3 N 14; Hirsch/Jacot-Guillarmod (n 27), pp. 160–161. Contra: Erard (n 10) (at least in the context of the HRA or the processing of health data); Thomas Probst, Die unbestimmte «Bestimmbarkeit» der von Daten betroffenen Person im Datenschutzrecht, AJP/PJA 10/2013, p. 1423: Probst seems to follow a nuanced approach: while criticizing the outcome of the Logistep case, he advocates for solutions where the viewpoint of the data recipients is taken into consideration in some circumstances. The Federal Data Protection Commissioner advocated for a more absolute approach: see FDPIC, 22e Activity Report 2014/2015, Bern 2015, p. 68.

31 See for instance Hirsch/Jacot-Guillarmod (n 27), who seem to focus on the Component 1; Meier (n 23), N 445 mentions: «*C'est par conséquent la théorie relative de la réidentification que l'on appliquera : seules les connaissances, moyens et possibilités, ainsi que l'intérêt propre de l'auteur (potentiel) du traitement doivent être pris en compte pour*

what Swiss and EU courts have done (despite the fact that scholars try to label such decisions as confirming either approach).[32] The situation may be summarized as follows:

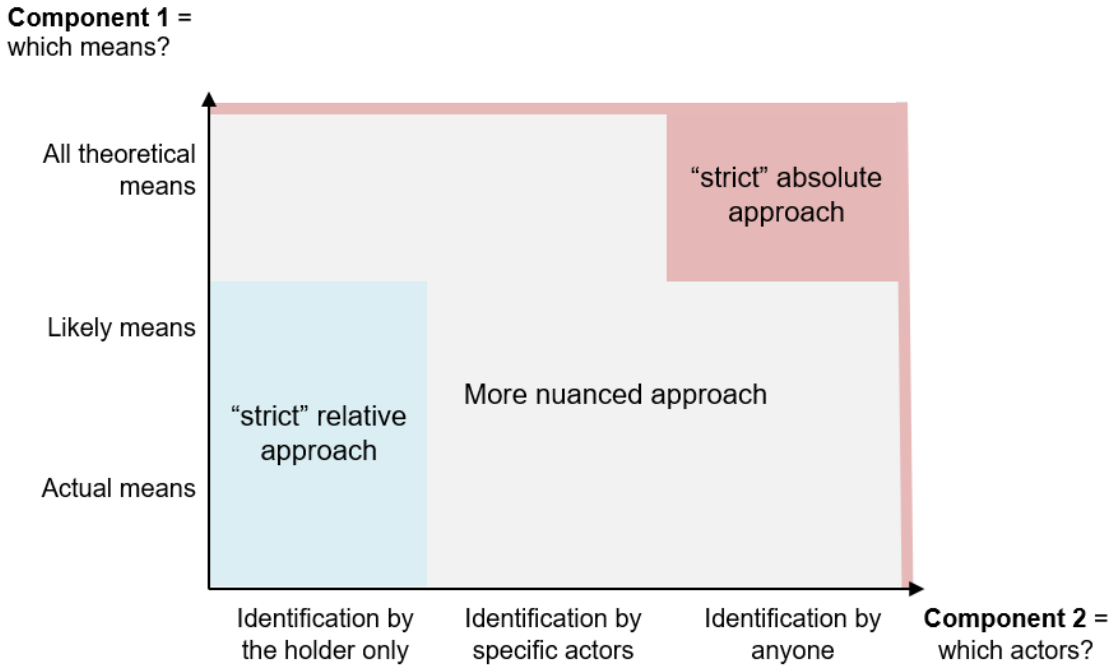

***Figure 1**: table of identifiability approaches. The smaller rectangle (in blue) shows the situations which fall into the «strict» relative approach, the larger rectangle in red encompasses all situations which fall into the «strict» absolute approach. All other zones are in-between situations that may be qualified as «nuanced» approaches.*

[19] Furthermore, the dichotomy «absolute» vs. «relative» is linked to concepts that are outdated, in particular regarding the roles of the actors as data controllers and joint data controllers (or controllers of the data file).[33]

[20] For the reasons indicated above, we advocate moving away from this dispute altogether. Each component of the identifiability should be assessed independently, taking into consideration (i) the rulings of the Swiss Federal Tribunal, which has on several occasions dealt with the concept of identifiability; (ii) the actual consequences each solution would entail; and (iii) the actual risks of (re-)identification, which require taking into account the data environment.

déterminer si la réidentification doit être envisagée [= components 1 and 2 cumulatively]. *Elle s'oppose à la théorie objective ou absolue, qui se satisfait déjà d'une possibilité théorique d'identification* [= component 1 only].

32 Those court cases are analyzed below. See *infra* chapters 3.5.2 and 5.

33 See *infra* chapter 3.5.2.

3.3. Assessing De-Identification and Re-Identification in the Data Environment

[21] Identification is linked to the concepts of data anonymization and data pseudonymization. Both refer to the action of processing personal data in such a manner that it can no longer be attributed to a specific data subject without the use of additional information.[34] Therefore, they affect the identifiability element of the personal data, by removing or masking the identifiers linked to the information (de-identification).[35]

[22] From a legal perspective, the difference between anonymization and pseudonymization is mainly that in case of anonymization, the reference to the person is irreversibly removed, while in case of pseudonymization it is only reversibly removed (a key, or assignment rule, being retained to enable re-identification).[36] In both cases, the output of the process results in data in which it is no longer possible to identify the data subjects (in case of pseudonymized data, except for the one holding the key). The information must have been sufficiently de-identified so that the data subjects may no longer be identifiable. While it is uncontroversial that anonymized data is no longer personal data,[37] the question is debated for pseudonymized data.[38]

[23] Anonymization (and pseudonymization from the viewpoint of the data recipient) is a heavily **context-dependent process** which requires consideration of data and its environment as a total system (i.e the data situation). **Data can only be determined as anonymized or not in relation to its environment**.[39] The risk of re-identification – which is never zero[40] – cannot be determined by examining the data alone; it does not reside solely in the properties of a dataset but rather arises from the interaction between that dataset, people, other data and the structures that shape those interactions such as security systems, governance practices, legislative frameworks, national policies on data sharing, etc. The core features – people, other data and structure – constitute the so-called **data environment**.[41]

---

34 See Art. 4(g) GDPR: pseudonymization is «*the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organizational measures to ensure that the personal data are not attributed to an identified or identifiable natural person.*».

35 We refer to de-identification to the actions of removing identifiers. Anonymization and pseudonymization, however, require not just a process but an ultimate «success state», in which the data cannot be attributed to an individual without the use of additional information. Thus, the data must not only be modified so that they are not directly identifiable, they must also be protected against re-identification. See Miranda Mourby et al., Are «pseudonymized» data always personal data? Implications of the GDPR for administrative data research in the UK, in: computer law & security review 34 (2018), pp. 222–233, p. 223.

36 For more information on those concepts, see: ENISA, 2021; Data Pseudonymisation: Advanced Techniques & Use Cases–Technical analysis of cybersecurity measures in data protection and privacy; Article 29 Data Protection Working Party, Opinion 05/2014 on Anonymization Techniques (adopted 10 April 2014).

37 Recital 26 GDPR; Zurich, HG190107O (n 13), para. 3.2.3 a).

38 *Infra* section 4.

39 Mark Elliot/Elaine Mackey/Kieron O'Hara, The Anonymization Decision-Making Framework: European Practitioners» Guide, 2n ed. Manchester, GB. UKAN 2020, p. 15.

40 Zero risk is not a realistic possibility if you are to produce useful data. Elliot/Mackey/O'Hara (n 39), p. 15.

41 Elliot/Mackey/O'Hara (n 39), p. 14 and 108.

Example 01

A public transport provider (PTP) collects personal data from its customers relating to public transport usage (= **data environment 1**). PTP wishes to share an anonymized version of the data with the local council of a city (LC), which wants to use it to help with infrastructure planning. PTP anonymizes the data by removing the direct identifiers, i.e. the customers' names and addresses, and by aggregating the detail on several key variables. However, it leaves in the dataset that will be shared with LC some key variables – which are of particular interest to LC – unchanged (= **data environment 2**).
By using a contract to stipulate how the data can be processed and accessed, PTP is placing controls on governance processes and infrastructure within environment 2 and thereby controls the disclosure risk associated with the data situation. The (anonymized) data within LC's environment is considered low risk even though it contains some detailed key variables. This is because the environment is restricted – few people have access to the data and their use of the data is clearly defined and managed.

Example 02

PTP wishes to make the same dataset available, but instead of sharing it with LC, it wishes to make it available in an open access environment (= **data environment 3**). In that less restricted environment, the data is unlikely to be considered «safe» (meaning that the risk of re-identification is too high), because no controls would be in operation.[42]

[24] This does not mean that examining the data itself is not important. Review of the data is necessary to determine a risk profile. Once a risk profile is developed for the data, the next step is to assess how the core features of the data environment (people, other data, and structure) may interact with the data to increase or negate the risk of re-identification.[43]

[25] The question of when data may be considered anonymous can thus be assessed by using the following formula: **the risk of re-identification depends on the content of the data and the context in which this data is processed**.[44] The data situation approach consists in assessing identifiability by considering the content of the data and the entire data environment.[45]

[26] It must be emphasized that the risk of re-identification must not be underestimated. This is particularly true for health data, which is closely linked to the physiological characteristics of a given person, and therefore more difficult to conceal, particularly in the era of precision medicine.[46] According to the case law of the Swiss Federal Tribunal, anyone who relies on anonymization or pseudonymization carried out by the recipient in order to prove the legiti-

42 This example is adapted from Elliot/Mackey/O'Hara (n 39), p. 31 ff.

43 Elliot/Mackey/O'Hara (n 39), p. 15.

44 Daniel Groos/Evert-Ben van Veen, Anonymized Data and the Rule of Law, in: European Data Protection Law Review Volume 6, Issue 4 (2020), p. 505; Elliot/Mackey/O'Hara (n 39), p. 15.

45 Groos/van Veen (n 44), p. 505.

46 See e.g. Ghislaine Issenhuth-Scharly, Autonomie individuelle et biobanques, Thesis, Geneva 2009, p. 212 ff.

macy of the disclosure of (original) personal data must assert and prove the effectiveness of the corresponding measures.[47] For this reason, relying on anonymization is always risky.

3.4. What Means Should Be Taken Into Consideration (*component 1*)?

[27] Under Swiss law, not every theoretical possibility of identification is sufficient for identifiability. If the effort is so great that, according to general life experience, it cannot be expected that an interested party will attempt to identify the data subject, there is no identifiability.[48] The question must be answered depending on the specific case, whereby in particular the possibilities offered by technology must also be taken into account.

[28] As stated by the Swiss Federal Tribunal, it is not only the level of effort objectively required to be able to assign a particular piece of information to a person that is relevant (**objective test**), but also what interest the person processing the data (or a third party) has in the identification (**subjective test**):

> *«What is important, however, is not only the effort that is objectively required to be able to assign a particular piece of information to a person, but also what interest the data processor or a third party has in identification.»*[49]

[29] Accordingly, not all eventualities of identification should be taken into consideration, but **only those likely to be used**, based on general life experience both from an objective and subjective standpoint. This view is also supported by the Federal Council in its Dispatch on the total revision of the FADP, which states that:

> *«The law does not apply to data that have been anonymized if re-identification by a third party is impossible (the data have been completely and definitively anonymized)* ***or only appears possible with such effort that no interested party will attempt it****. This last rule also applies to pseudonymized data.»*[50]

[30] EU law recognizes the same approach possibilities, noting right from recital 26 of the GDPR:

> *«To determine whether a natural person is identifiable, account should be taken of* ***all the means reasonably likely to be used****, [taking into account] all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments.»*

[31] The Breyer case of the CJEU, discussed below,[51] also follows this approach.

---

47 Judgement of the Federal Tribunal 4A_365/2018 of 26 February 2018, para. 5.2.2.

48 TF 4A_365/2017 (n 23), para. 5.

49 TF 4A_365/2017 (n 23), para. 5 (unofficial translation of «Von Bedeutung ist indessen nicht nur, welcher Aufwand objektiv erforderlich ist, um eine bestimmte Information einer Person zuordnen zu können, sondern auch, welches Interesse **der Datenbearbeiter oder ein Dritter** an der Identifizierung hat»); see also ATF 138 II 346 para. 6.1; ATF 136 II 508 para. 3.2; Zurich, HG190107O (n 13), para. 3.2.3 a).

50 Federal Council Dispatch revFADP (n 22), 6640 (unofficial translation of: «La loi ne s'applique pas aux données qui ont été anonymisées si une ré-identification par un tiers est impossible (les données ont été anonymisées complètement et définitivement) **ou ne paraît possible qu'au prix d'efforts tels qu'aucun intéressé ne s'y attèlera**. Cette dernière règle vaut aussi pour les données pseudonymisées»).

51 *Infra* chapter 5. See also Judgment of the Tribunal of the European Union T384/20 of 4 May 2022, N 68.

[32] Hence, neither Swiss law nor EU law follows an absolute approach in relation to the component 1 of identifiability. Only the means **likely to be used** should be taken into consideration to assess identifiability.

3.5. Which Actors Should Be Taken Into Consideration (*component 2*)?

3.5.1. In General

[33] The «absolute vs. relative» debate remains too schematic even when focusing only on the second component, i.e. determining the relevant actors for the identification test. There are wide *nuanced* possibilities, ranging from taking into account only the viewpoint of the data holder to taking into account the viewpoint of any person.

[34] The «strict absolute approach» (i.e. that identification by anyone is sufficient) must be rejected. It would lead, in practice, to the classification as personal data of all kinds of information, no matter how insufficient such information may be in and of itself to facilitate the identification of a data subject. It would never be possible to rule out, with absolute certainty, the possibility that there may be a third party in possession with additional data revealing a person's identity. Data holders – unaware of the fact that the data they process qualifies as personal data – would be unable to comply with their obligations.[52]

[35] Conversely, the «strict relative approach» (i.e. that information is only personal data if the holder of the information can identify the individual) must also be rejected. First of all, because it would lead to inappropriate outcomes:

Example 03

A newspaper receives photographs about two young adults having intimate relations in a public place. The newspaper cannot identify the individuals, but they are (easily) recognizable in the photographs for the persons who know them. Under a strict relative approach, the photographs would not be personal data for the newspaper and thus their publication (unblurred) would not constitute a processing of personal data within the meaning of the FADP.[53]

[36] Moreover, the «strict relative approach» contradicts the positions repeatedly held by the Swiss Federal Tribunal as will be analyzed below.

3.5.2. The Logistep case

[37] The «Logistep» case[54] concerns a company, Logistep AG, which developed a software to monitor various peer-to-peer networks for infringing copyrighted works. The software collected and stored certain pieces of information, in particular the dynamic IP addresses through which

52 Case C582/14 of the CJUE, Opinion of Advocate General Campos Sánchez-Bordona, 12 May 2016, N 65.

53 See also the example cited by the Swiss Federal Tribunal in ATF 136 II 508para. 3.4. Probst(n 30), p. 1423, considers that in that case, there is a rebuttable presumption that the data subjects are identifiable.

54 ATF 136 II 508.

the illegal downloads were made. The data collected this way was then passed on to the copyright holders of the infringed works and used by them to identify the concerned individuals. To do so, they had to file criminal complaints against unknown persons to obtain the identity data within the framework of the right to inspect files.[55] This data was then used to assert claims for damages. Logistep argued that it did not process personal data because it had no means of identifying the persons behind the connections. The Swiss Federal Tribunal ruled that:

> *Whether information can be linked to a person on the basis of additional data, i.e. whether the information relates to an identifiable person (Art. 3 lit. a FADP),* ***is assessed from the perspective of the respective holder of the information*** *[. . .].*
>
> ***In the case of disclosure of information, it is sufficient if the recipient is able to identify the data subject.*** *[. . .] This means for the present case that it is not a prerequisite that the copyright infringers are already identifiable for the respondent. Rather, it is sufficient if they become so for the copyright holders after the relevant data has been handed over. If this is the case (see below), the FADP also applies to the [data holder] itself. To decide otherwise would mean to apply the FADP only to the individual recipients, but not to the person who obtains the data in question and disseminates them. This would be contrary to the purpose of the Act.*[56]

[38] Accordingly, the Swiss Federal Tribunal considered that whether information can be associated with a person on the basis of additional details, i.e. whether the information relates to an identifiable person, **is assessed from the perspective of the respective holder of the information** (data holder). However, in case of disclosure of the information, identifiability must also be **assessed from the perspective of the recipient of the information (data recipient)**.[57] We agree with this position. As stated above, the risk of identification depends not only on the data holder, but more generally on the data situation, which requires to consider the entire environment, including in particular anticipated data recipients.[58]

[39] The Swiss Federal Tribunal has confirmed its position in a subsequent published decision regarding Google Street View:

> *«Whether information can be associated with a person on the basis of additional details, i.e. whether the information relates to an identifiable person (Art. 3 lit. a FADP), is assessed*

55 This is because the owners of dynamic IP addresses can usually only be identified with the help of the internet service providers (ISP) which assigned the address. The ISP is obliged to maintain secrecy vis-à-vis third parties (Art. 43 of the Telecommunications Act [RS 784.10]).

56 ATF 136 II 508 para. 3.4 (unofficial translation of: «Ob eine Information aufgrund zusätzlicher Angaben mit einer Person in Verbindung gebracht werden kann, sich die Information mithin auf eine bestimmbare Person bezieht (Art. 3 lit. a DSG), beurteilt sich aus der Sicht des jeweiligen Inhabers der Information [. . .] Im Falle der Weitergabe von Informationen ist dabei ausreichend, wenn der Empfänger die betroffene Person zu identifizieren vermag. [. . .] Dies bedeutet für den vorliegenden Fall, dass nicht vorausgesetzt ist, dass die Urheberrechtsverletzer bereits für die Beschwerdegegnerin bestimmbar sind. Vielmehr genügt es, wenn sie es nach Übergabe der entsprechenden Daten für die Urheberrechteinhaber werden. Trifft dies zu (dazu sogleich), so gelangt das Datenschutzgesetz indessen auch auf die Beschwerdegegnerin selbst zur Anwendung. Anders zu entscheiden würde bedeuten, das Datenschutzgesetz nur auf die einzelnen Empfänger anzuwenden, nicht aber auf die Person, welche die betreffenden Daten beschafft und sie verbreitet. Dies würde dem Zweck des Gesetzes zuwiderlaufen.»).

57 ATF 138 II 346 para. 6.1; ATF 136 II 508 para. 3.4; Zurich, HG190107O (n 13), para. 3.2.3 a). This is, however, debated.

58 *Supra* chapter 3.3.

> *from the perspective of the respective owner of the information. In the case of disclosure of information, it is sufficient if the recipient is able to identify the data subject*».[59]

[40] Hence, if the intended recipients have the capacity to identify the individual (as the case may be through the use of additional information), the data will be personal data from the point of view of both the data provider (Logistep) and the data recipients (the right holders).

[41] Identifiability is given even if the information that enables the identification of the person concerned must be requested – either by the data holder or the data recipient – in a procedure with the involvement of a state authority.[60] From the point of view of the data holder, the fact that a third party (i.e. the data recipient) must take specific action (i.e. file a criminal complaint) to identify the data subjects is therefore irrelevant, unless it cannot be expected (according to general life experience) that the data holder or the data recipient will attempt to identify the data subject because the overall effort required would be excessive.[61]

> «*[...] the necessity of an action by a third party is irrelevant as long as the overall effort of the client to identify the person concerned is not so great that, according to general life experience, it could no longer be expected that the client would carry it out (cf. E. 3.1 above). This is to be assessed against the background of the concrete circumstances of the individual case. [...].*»[62]

[42] Accordingly, it is not relevant that the data holder does not hold the additional information required to identify the individuals, insofar as the persons to whom it hands over this data have the means to re-identify the persons concerned. It is also not relevant that the data recipients need to take extra steps to access the additional information required to identify the individuals (even if these means required quite a lot of effort, provided it is likely that they would go to the trouble).[63]

[43] Scholars have commented on – and often criticized – this ruling as establishing an absolute conception of personal data.[64] Others have considered that this ruling confirmed that Swiss law

---

59 ATF 138 II 346 para. 6.1 (unofficial translation of «Ob eine Information aufgrund zusätzlicher Angaben mit einer Person in Verbindung gebracht werden kann, sich die Information mithin auf eine bestimmbare Person bezieht (Art. 3 lit. a DSG), beurteilt sich **aus der Sicht des jeweiligen Inhabers der Information**. **Im Falle der Weitergabe von Informationen ist dabei ausreichend, wenn der Empfänger die betroffene Person zu identifizieren vermag**»); the Zurich Handelsgericht has endorsed this position in its judgment HG190107O (n 13), para. 3.2.3 a).

60 ATF 136 II 508 consid. 3.5. This position has been endorsed by the Appellate Court of the Basel-State Canton (see: Basel-Stadt ZB.2019.3 (n 14), para. 4.2.1) and the Handelsgericht of Zurich (see: Zurich, HG190107O (n 13), para 3.2.3 a).

61 ATF 136 II 508 para. 3.5.

62 ATF 136 II 508 para. 3.5 (unofficial translation of: «[...] die Notwendigkeit des Tätigwerdens eines Dritten so lange unmassgeblich ist, als insgesamt der Aufwand des Auftraggebers für die Bestimmung der betroffenen Person nicht derart gross ist, dass nach der allgemeinen Lebenserfahrung nicht mehr damit gerechnet werden könnte, dieser werde ihn auf sich nehmen (vgl. E. 3.1 hiervor). Solches ist vor dem Hintergrund der konkreten Umstände des Einzelfalls zu beurteilen [...].)».

63 In the Logistep case, the information was difficult to obtain: it required filing a criminal complaint, linking the dynamic IP address to a connection, consulting the file, verifying who was the person actually behind the computer, etc. However, the copyright holders (the data recipients) specifically collected the data from Logistep for the purpose of identifying the individuals.

64 For the excessive approach, see e.g. Philippe Meier, Federal Data Protection and Information Commissioner v. Logistep AG (appeal in civil matters), 8 September 2010, JdT 2011 II 446, p. 462 ff. For a limitation of this judgment to the facts of the case, see in particular: Hirsch/Jacot-Guillarmod (n 27), p. 161 and more precisely the references cited in note 89.

follows a relative approach.[65] Neither view is entirely correct.[66] The Swiss Federal Tribunal considered the means likely available to the data holder and the intended recipients of the data holder (and not anyone else). In this case, the recipients received the data specifically for the purpose of identifying the concerned individuals. It was clear that they had the possibility and willingness to identify the data subjects. **Accordingly, the Swiss Federal Tribunal followed a nuanced approach, in which only the likely means available to the data holders and the intended recipients must be taken into consideration to assess identifiability**. The position of the Swiss Federal Court is aligned with the latest developments in (re)identifiability assessment pursuant to which the entire data environment must be considered (including in particular the data recipients).[67]

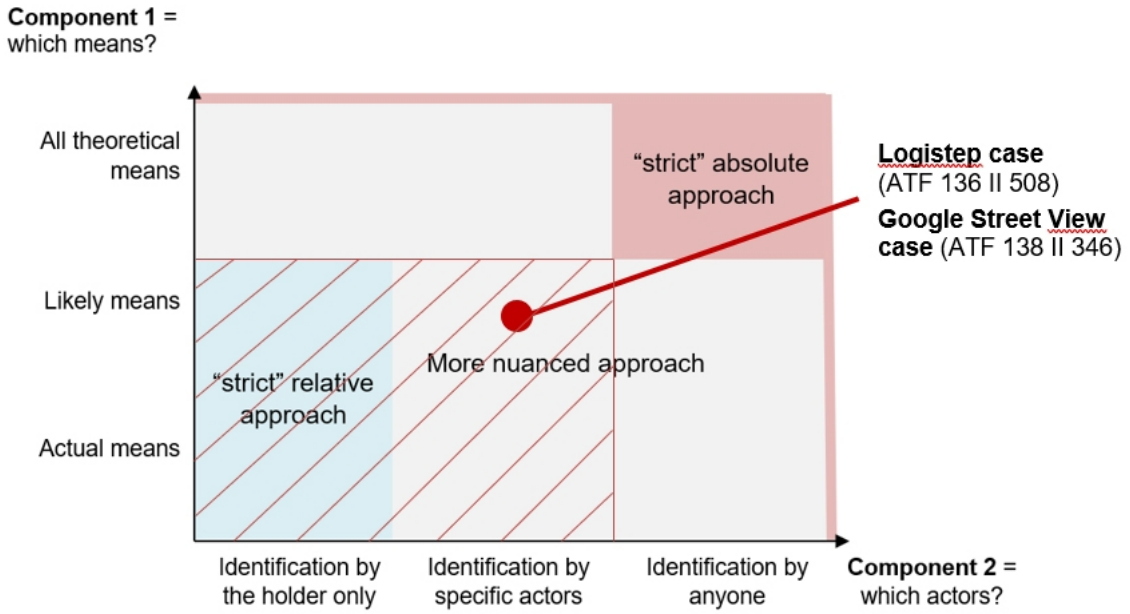

***Figure 2**: positioning the Swiss Federal Tribunal cases on «table of identifiability approaches» of Figure 1. The striped area shows all situations which would fall under the identifiability test of the Swiss Federal Tribunal.*

[44] We stress that, in the Logistep case, the Swiss Federal Tribunal did not address the situation where a data recipient receives data which is anonymous for the recipient but personal data for

65 Rosenthal, Das neue Datenschutzgesetz (n 30), N 19 who refers to «*the relative approach confirmed by the Swiss Federal Tribunal*» and cites the Logistep case (ATF 136 II 508, para. 3.2). We note that the paragraph cited by this author refers to the first component of the approach (which means should be taken into consideration), rather than the second component (whose point of view). On the second component, the paragraph states that one must take into account the interest of the data holder or a third party («*welches Interesse der Datenbearbeiter oder ein Dritter an der Identifizierung hat*»), which does not seem compatible with a «strict relative approach».

66 These contradictions are an inherent consequence of the flaws the concepts of «absolute» and «relative» suffer. See *supra* chapter 3.2.

67 *Supra* chapter 3.3.

the data provider (e.g. in case of sharing of pseudonymized data).[68] We will analyze this second situation below.[69]

[45] Finally, we note that the characterisation of the data as personal data or not from the viewpoint of Logistep should not have had material consequences on its obligations. In that case, Logistep must be considered as a **joint data controller**,[70] together with the rights holders, in relation to the collection and communication of the dynamic IP addresses. Indeed, Logistep developed and operated the software capable of collecting such data. The right holders, on their end, decided which copyright works should be monitored, and thus participated in the decision on the purposes and the means of the processing activities. To be considered as a joint controller, it is not necessary to have actual access to the data: the fact that one of the parties does not hold any data, or only holds data which it cannot link to an individual (e.g. coded data) does not impact its qualification as joint data controller.[71] A data holder in a situation similar to the one of Logistep will assume obligations under the FADP independently from the qualification of the data as personal data or not, because of its role as joint data controller.

Example 04

A website operator places a Facebook plug-in (the «Like» button) on its website, which permits Facebook to collect data from the website's visitors. The website operator does not have access to the personal data collected this way. However, it acts as a joint controller with Facebook with regard to the collection and transmission of the data.
As a result, it needs to comply with the obligations of a data controller in relation to these activities. Whether the data qualifies (or does not qualify) as personal data for the website operator does not impact its obligations.[72]

3.6. Interim Conclusion

[46] It results from the case law of the Swiss Federal Tribunal that **in case of data sharing, the identifiability test must take into account the relevant data environment from the point of view of the data holder**. This requires an assessment of whether identifiability is reasonably likely in such data environment. The same information can be personal data to one organization, but anonymous information in the hands of another organization.

[47] In accordance with the case law of the Swiss Federal Tribunal, we consider that **from the point of view of the data holder (Company ABC)**:

> (a) Information will be considered as personal data if Company ABC can identify the individuals, even if this requires the assistance of a third party (provided that it remains likely that those means will be used).

68 Certain authors seem to consider that the decision could also apply to this situation. See e.g. Erard (n 10), p. 611.

69 *Infra* chapter 6.

70 Art. 5 let. j et 33 nLPD. Art. 26 GDPR.

71 See Jotterand/Erard (n 6), N 82 and the references cited therein.

72 This example corresponds to the CJEU Case C-40/17of 29 September 2019, which is transposable under Swiss law. See Célian Hirsch, Fashion ID, Facebook, le bouton «j'aime» et la notion de coresponsable du traitement, in: www.lawinside.ch/805/.

(b) The information will also be considered as personal data even if Company ABC cannot directly identify the individuals, but shares the data with third parties who can (and should be expected to).[73]

3.7. Restating the Identifiability Test

[48] In accordance with the principles stated above, we propose to restate the approach of identifiability as follows:

- Identifiability must be assessed from the viewpoint of each data holder. Personal data is a dynamic concept. It can be personal data for one person and anonymous data for another. This is aligned both with all decisions from Swiss courts, and the doctrine on this subject.[74]
- The identifiability test must only consider the means reasonably likely to be used and not every theoretical possibility of identification.[75] This entails:
    - an objective assessment: the level of effort objectively required to be able to assign a particular piece of information to a person; and
    - a subjective assessment: what interest the relevant person (see below) has in the identification.
- It is not relevant that someone, somewhere, holds the means necessary (and would be willing) to identify the individual. The assessment must be based on the entire data environment.[76] In simple cases, the data environment is limited to the data holder (one organization), in which case the identifiability test is based on its viewpoint:
    - If the data holder controls the additional information required for identification (e.g. the identification key of a pseudonymized dataset), then the data is personal data for the data holder;
    - If the data holder does not control the additional information, but such information exists and it is likely that the data holder will use it (following the test under let. (b) above), then the data is personal data for the data holder[77]; and
    - If the additional information necessary to identify the individuals does not exist, or it exists but it cannot be expected that the data holder will attempt to identify the data subject, there is no identifiability and the data is not personal data for the data holder.
- When a data holder (Company ABC) intends to share the data with third parties, then the data environment changes and the identifiability assessment – to determine if the data is

73 Probst (n 30), p. 1423 considers that this should be the case only in exceptional cases, namely if the data provider knew or should have known about the means available to the data recipient and nevertheless failed to take the necessary measures to prevent the identification. In case of public dissemination, there would be a rebuttable presumption that the data subjects are identifiable.

74 *Supra* chapter 3.5.

75 *Supra* chapter 3.4.

76 *Supra* chapter 3.3.

77 It is irrelevant if the data holder does not have access to the additional information itself.

personal data from the viewpoint of Company ABC – must include the means reasonably likely to be used by the data recipients.[78]

- For a data recipient, the fact that the data was personal data for the previous data holder is irrelevant. Thus, one must assess if the data recipient – which is a new data holder – is likely to identify individuals, using the same test as above.

4. Handling and Transferring Pseudonymized or Coded Data

[49] The rules applicable in the event of the transfer of pseudonymized data to a third party who does not have access to the re-identification key is controversial in Switzerland, in particular as to whether pseudonymized data should be considered as anonymous data from the point of view of such recipient.[79] The controversy is directly linked to the debate on the «*absolute*» *vs.* «*relative*» *approach to identifiability*, which as we have seen must be abandoned.[80]

4.1. Under the FADP

[50] The FADP does not specifically address the issue of the transfer of pseudonymized data to third parties and the Swiss Federal Tribunal has not yet rendered any decision directly addressing this issue. In the Logistep case, it addressed a reversed situation (where the provider cannot identify the individuals, but the recipient can). The court only stated that «*In the case of disclosure of information, it is sufficient if the recipient is able to identify the data subject*».[81] In our opinion, this decision does not mean that, as a general rule, a piece of information must be considered as personal data if one of the two parties to the data exchange is able to (re-)identify the data subject.[82]

[51] The Commercial Court of Zurich ruled on this topic in two separate instances (in 2017 and 2021) that pertained to the transfer by a bank of a pseudonymized list of customers (i.e. data for which the bank kept a concordance table) to the U.S. Department of Justice. Both cases thus related to a situation in which a party acting as **data controller** wished to **transfer pseudonymized data** (which is personal data for that party, because it retained the key) to a third party acting as **independent data controller**. The court ruled that:

- For all those who have access to the key, pseudonymized personal data remain personal data within the meaning of the FADP.

78 See the Example 01 and Example 02 above. We stress that how you conduct this assessment will vary depending on the roles assumed by the data provider and the data recipient(s).

79 Erard (n 10), p. 609.

80 *Supra* chapter 3.2.

81 ATF 136 II 508 para. 3.4 (unofficial translation of: «*Im Falle der Weitergabe von Informationen ist dabei ausreichend, wenn der Empfänger die betroffene Person zu identifizieren* vermag.»).

82 Contra: Erard (n 10), 611.

- For persons who do not have access to the key and do not have other knowledge to be able to link the data to a specific person again, pseudonymized personal data, on the other hand, no longer constitute personal data [...].[83]

[52] The Zurich court, however, stressed that to exclude re-identification, it is often not sufficient to remove clearly identifying characteristics such as first name, last name, date of birth and address. The technologies available today, particularly those linked to *Big Data*, leave little room for irreversible anonymization (or equivalent pseudonymization). As a result, the court considered in both cases that the bank had not proven that it had taken sufficient steps to prevent re-identification by the U.S. authorities.

[53] In the first case (from 2017), the bank appealed to the Swiss Federal Tribunal, which rejected the appeal without developing any particular arguments pertaining to the matter at hand, on the basis that the bank did not sufficiently demonstrate the effectiveness of the pseudonymization measures.[84] Certain authors considered that the decision of the Federal Court rejecting the appeal against the decision of the Zurich Handelsgericht provided a welcome clarification and validated the so-called relative approach, even putting an end to the uncertainties associated with the Logistep decision.[85] We do not believe this to be the case. The Swiss Federal Court did not address the question of identifiability in its decision, nor the issue of absolute versus relative approaches.[86] Furthermore, both the Zurich court and the Swiss Federal Tribunal analyzed whether the data recipient, i.e. the DoJ, could use additional knowledge to identify the individuals and assess the responsibility of the bank on this basis,[87] and considered that if that was the case, the data would be considered personal data for the DoJ and for the Bank. Accordingly, the obligations of the bank to comply with the FADP depended on the possibility for the data recipient to (re-)identify the dataset, thus taking into consideration the entire data situation. This is, in our opinion, a confirmation of the previous published decisions of the Swiss Federal Tribunal (ATF 136 II 508 and ATF 138 II 346).

83 Zurich, HG190107O (n 13) (unofficial translation of: «*Für alle, die Zugang zum Schlüssel haben, bleiben pseudonymisierte Personendaten weiterhin Personendaten im Sinne des DSG. Für Personen, die keinen Zugang zum Schlüssel haben und auch nicht über andere Kenntnisse verfügen, um die Daten wieder einer bestimmten Person zuordnen zu können, stellen pseudonymisierte Personendaten hingegen keine Personendaten mehr dar*»).

84 TF, 4A_365/2017 (n 23), para. 5.2.2.

85 Hirsch/Jacot-Guillarmod (n 27), p. 162.

86 TF, 4A_365/2017 (n 23), para. 5.2.2.

87 TF, 4A_365/2017 (n 23), para. 3.5.

Example 05

In the situations considered by the Zurich Courts, the bank wished to transfer pseudonymized datasets (list of customers) to the DoJ. The datasets were (rightfully) considered pseudonymized because the bank kept the key necessary to re-identify the customers.
If we consider, for the purpose of the discussion, that the bank had not retained this key, then the dataset would have been considered anonymous from the point of view of the bank. In our opinion, prior to sending this now «*anonymous*» list to the DoJ, the bank would still have had to assess whether the DoJ is able (and willing) to re-identify the data subjects. If this were found to be the case, the dataset would – in the context of this transfer, from the point of view of the bank – still have been considered as personal data.[88]

[54] The Zurich Court further considered that if personal data is anonymized or pseudonymized before being disclosed abroad in such a way that its recipient abroad can no longer establish a personal reference, there is no cross-border disclosure of personal data within the meaning of Art. 6 FADP (= Art. 16–18 revFADP).[89] Although the Swiss Federal Tribunal neither confirmed nor denied this assertion, we consider that this position proves our point further. In those specific circumstances, the personal data is never made available abroad and the protection granted by the FADP is not necessary.[90]

[55] This does not mean, however, that no rule applies to the transfer of pseudonymized data. From the point of view of the data provider (the original data holder), the data remains personal data. As such, **both the pseudonymization process and the communication of the data to a third party remain processing activities carried out on personal data**[91] **which – if they breach the privacy of the individuals concerned**[92] **– must rely on a justification** (Art. 12 para. 1 and 13 FADP/30 and 31 FADP).[93] When assessing if those processing activities constitute a breach of privacy, and if so, whether the breach is justified, it is necessary to take into account the use that will be made of the data by the data recipient.[94]

[56] The current positions under Swiss law may thus be summarized as follows:

88 This means that identification (and the risk of re-identification) must be assessed based on the entire data situation, which include the intended recipients. Considering that in the matter at hand, the data is not personal data for the bank and can freely be sent to the DoJ would have an unjustified impact on the personality right of the individuals.

89 Zurich, HG190107O (n 13), para. 3.2.3 a).

90 Art. 6 FADP/16–18 revFADP aims at ensuring that the privacy of the data subject will not be endangered by the absence of an appropriate legislation.

91 The Swiss Federal Tribunal confirmed that anonymization and pseudonymization constitute a processing of personal data within the meaning of Art. 3 lit. e FADP, even if the result of the action is no longer personal data: TF, 4A_365/2017 (n 23), para. 5.2.2.

92 «*Atteinte à la personnalité*»; «*Verletzung der Persönlichkeit*». More specifically a breach of the individuals' right to informational self-determination («autodétermination informationnelle»). Meier (n 23), 1531.

93 TF, 4A_365/2017 (n 23), para. 5.2.2.

94 In general it is likely that applying anonymization or pseudonmization techniques to personal data will not breach the personality rights of the individuals, or if so will be justified. But this should be assessed on a case-by-case basis. In connection with the HRA, the Federal Council in its Dispatch has considered that the anonymization of non-genetic personal health data for research purposes is unconditionally permitted (Federal Council Dispatch of 21 October 2009 regarding the Human Research Act (HRA), FF 2009 7259, p. 7338).

- Pseudonymized data is personal data for all those which have access to the key or those which nevertheless can identify the individuals (based on means likely to be used),
- For the recipients who do not have access to the key and cannot otherwise identify the individuals, the situation remains debated, although the prevailing position is that such data should not be considered as personal data for the data recipient.
- For the data holder, the data remains personal data and both the pseudonymization process and the communication of the data to a third party constitute processing activities carried out on personal data which – if they breach the privacy of the individuals concerned – must rely on a justification. The data provider must ensure that it complies with its obligations under the FADP.

[57] We stress that considering that the data is not personal data for the data recipient which receives pseudonymized data without the corresponding key may actually lead to additional obligations on the data provider. Indeed, in that case, in order to ensure that it will be able to comply with its own obligations under the FADP, the data provider must impose appropriate obligations. Accordingly, the same data set will be considered personal data for the data provider, but anonymous data for the data recipient. The actual consequences of this situation must be analyzed on a case-by-case basis.[95]

Example 06

Company ABC stores encrypted data about its customers on the servers of a cloud service provider (CSP). The CSP has no access to the key (and no likely mean to decrypt the data). A customer of Company ABC contacts Company ABC and asks for the deletion of outdated information. Company ABC cannot refuse to act on the basis that the data is stored with a third party for whom the data is not personal data. Company ABC must secure its ability to instruct the CSP on the use of the data even if the data is not personal data for the service provider.[96]

95 See *Infra* chapter 4.3.

96 It is customary in cloud service agreements to have different provisions for the handling of personal and non-personal data. These agreements will need to be interpreted when it is agreed that the data is only personal data for one of the contractual parties.

Example 07

Continuing on the same example: a security breach occurs at that CSP, resulting in Company ABC's data being disclosed on the dark web.
Under the revised FADP, this constitutes a data breach triggering the consequences of Article 24 revFADP. Company ABC will thus have to assess if a notification is required. If we consider that the data is not personal data for the CSP, then the CSP will have no statutory obligation to notify the breach to Company ABC (Art. 24 para. 3 revFADP will not apply). Because the obligations on Company ABC do not depend on the qualification of the data from the point of view of the CSP, Company ABC should secure through appropriate contractual provisions that the CSP will notify to Company ABC all data breaches (even if there is no personal data from the point of view of the CSP).

Example 08

Company ABC asks its customers to consent to having their data used for the purpose of a research conducted by several insurance companies, which all share the data of their customers in a pseudonymized manner. The data is appropriately pseudonymized and the other insurances companies cannot identify the customers of Company ABC.
A customer of Company ABC contacts Company ABC and withdraws her consent to have her data used as part of the research. Company ABC must act upon the withdrawal and cannot allow the other insurance companies to continue using the data on the basis that it would not be personal data from their point of view.

[58] For the same reasons, even in case the data is deemed as anonymous from the point of view of the data recipient, the data provider must ensure that the data recipient appropriately protects the data against security incidents. Further, it must ensure that the circle of data recipients is adequately limited. Otherwise, the risk of re-identification may be so important that the data will no longer be considered as sufficiently anonymized.

Example 09

A telecommunication provider Company (TelCo) shares a dataset containing pseudonymized information concerning its clients (individuals) with a researcher who wishes to conduct a research project. The parties assume that the FADP does not apply because the researcher cannot identify the data subjects. After publishing the result of her research, the researcher is criticized by the scientific community about the methodology of her research. To demonstrate the soundness of her work, she publishes online the entirety of the (still pseudonymized) dataset. As it happens, an attacker can easily identify individuals from this dataset by recouping it with additional information. This constitutes a personal data breach pursuant to the FADP. TelCo must ensure in its contract with the researcher that this situation is appropriately covered.

4.2. Under the revFADP

[59] The above considerations also apply under the revFADP. The revised Act does not introduce any novelty pertaining the issue at stake. The Dispatch of the Federal Council on the total revision of the FADP does not directly address this issue, except for the following citation:

> «*The law does not apply to data that has been anonymized if re-identification by a third party is impossible (the data has been completely and definitively anonymized) or only appears possible with such effort that no interested party will attempt it. This last rule also applies to pseudonymized data.*»[97]

[60] There is therefore no reason to deviate from the considerations above applicable to the current FADP.

4.3. Interim Conclusion

[61] In our opinion, the legal regime pertaining to the handling and transferring of pseudonymized (or coded) data should be based on the core principle that **what is not permitted for the data controller should not become permissible only because it transfers pseudonymized or coded data to a third party which does not have the identification key**. Accordingly, what a data controller cannot do, it cannot authorize (or permit) a third party to do.[98]

Example 10

Any organization must protect personal data – including pseudonymized data – through adequate technical and organizational measures (Art. 7 FADP). If such data is (coded) health-related personal data used in the context of medical research, the organization must in addition comply with the requirements of Art. 43 HRA and Art. 5 HRO.
The organization cannot circumvent these obligations by outsourcing the pseudonymized/coded data to a third party. The data should be afforded the same level of security if storage is undertaken directly by the organization or outsourced to a third party.

97 Unofficial translation: «La loi ne s'applique pas aux données qui ont été anonymisées si une ré-identification par un tiers est impossible (les données ont été anonymisées complètement et définitivement) ou ne paraît possible qu'au prix d'efforts tels qu'aucun intéressé ne s'y attèlera. **Cette dernière règle vaut aussi pour les données pseudonymisées**»; Federal Council Dispatch revFADP (n 22), 6640.

98 This principle is aligned with the requirement of Art. 10a para. 1 let. a FADP (= Art. 9 para. 1 let. a revFADP), which provides that the processing may be assigned to a third party if «*the data is processed only in the manner permitted for the instructing party itself*».

Example 11

A company holds a database containing personal data of its individual clients having opted-out to the processing of their data for market analysis purposes. Accordingly, the company is not authorized to process such data for market analysis purposes (Art. 12 para. 2 let. b FADP).
The company cannot share the dataset in a pseudonymized manner to a third party and ask it to perform the market analysis operations and deliver the result back to the company. The fact that the data is or is not personal data for the third party is irrelevant (from the point of view of the company).

[62] There are various means to ensure compliance with this key requirement. The first is to consider that, in all these situations, the pseudonymized data must be considered as personal data for the data recipient. But doing so is not necessarily required, as long as the data provider is held accountable for the processing operations carried out by the data recipient(s) in the relevant data environment.

[63] This requires distinguishing between situations which are fundamentally different. As an example, the sharing of pseudonymized data with a completely independent data controller, as was the case of the bank sharing data with the U.S. Department of Justice, is different from the situation of a company sharing pseudonymized data with a cloud service provider. In the first case, it makes no difference that the bank kept a key (cross-reference table), whereas in the second case it is very relevant.

[64] Furthermore, one must distinguish for each obligation under the FADP if such obligation remains relevant in the context of the sharing of pseudonymized data. For instance, the obligations to comply with the FADP general principles[99] and to have a lawful basis in case the personality rights of the individuals are breached,[100] remain applicable for the original data holder since the data is still personal data for it. However, because there is no disclosure of personal data (what the recipient receives is not personal data, even if it is so for the data provider), the rules limiting the disclosure of sensitive and/or personal data,[101] or those restricting cross-border transfers,[102] will not apply.

5. Considerations under the GDPR

[65] Pursuant to the GDPR, data is personal when the *controller or another person* is able to identify the data subject by using «*means reasonably likely to be used*». Pseudonymization is generally not considered as a method of anonymization, but as a security measure reducing the linkability of a dataset with the original identity of a data subject.[103] The most obvious benefit of

99 Art. 4 and 5 FADP; 6 revFADP.

100 Art. 13 and 13 FADP; Art. 30–31 revFADP.

101 Art. 12 para. 2 let. c/ 19 FADP; Art. 30 para. 2 let. c/ 36 revFADP.

102 Art. 7 FADP; 16 ff. revFADP.

103 Recitals 26 and 28 GDPR; Article 29 Data Protection Working Party. Opinion 05/2014 on anonymisation techniques. European Commission, p. 20.

pseudonymisation is to hide the identity of the data subjects from any third party (other than the pseudonymising entity, i.e. the entity responsible for pseudonymisation).[104] As shown by documents issued during the preparation phase of the GDPR, the final wording of the key provisions of the GDPR for this issue (definitions of personal data, pseudonymized data, and the recitals 28 and 29 GDPR) is the result of a compromise between actors advocating for a more absolute approach and those advocating for a more relative approach of the concept of personal data.[105]

[66] In the Breyer case, the Court of Justice of the European Union (CJEU) had to interpret the Directive 95/46 (precursor to the GDPR) to determine whether a dynamic IP address and data relating to the date and time of connection constituted personal data in the eyes of an online media service provider, on the understanding that such data did not allow for identification without cross-checking against data held by the internet service provider. In a nutshell, the court considered that it was necessary to determine «*whether the possibility to combine a dynamic IP address with the additional data held by the internet service provider constitutes a means likely reasonably to be used to identify the data subject*.» It concluded that the data was personal, but only because of the existence of legal channels enabling the competent authority to obtain identifying information from the internet service provider in the event of a cyber-attack. In the absence of these channels, the data would not have been considered personal simply because a known third party could identify them:

> «*Thus it appears that the online media services provider has the means which may likely reasonably be used in order to identify the data subject [. . .] a dynamic IP address registered by an online media service provider [. . .] constitutes personal data [. . .] in relation to that provider, where the latter has the legal means which enable it to identify the data subject with additional data which the internet service provider has about that person.*»[106]

[67] According to the Breyer case, data (such as dynamic IP addresses) is personal data if the data holder (in that case the website operator) has «legal means» enabling the identification of the person associated with the data with the help of third parties (in that case, the internet service provider). This was considered to be the case in Breyer. The Breyer case may be – and has been – interpreted in many different ways. To us, Breyer indicates that the CJEU excluded a «strict absolute approach», because dynamic IP addresses would be considered as personal data for everyone since they are for the ISP, independently from the actual likely means to access that information.

104 ENISA (n 36), p. 10.

105 Julien Rossi, Protection des données personnelles et droit à la vie privée: enquête sur la notion controversée de «donnée à caractère personnel», Science politique. Université de Technologie de Compiègne, 2020 (Thesis), p. 549.

106 Case C-582/14 Patrick Breyer v Bundesrepublik Deutschland [2016] ECLI: EU: C: 2016:779.

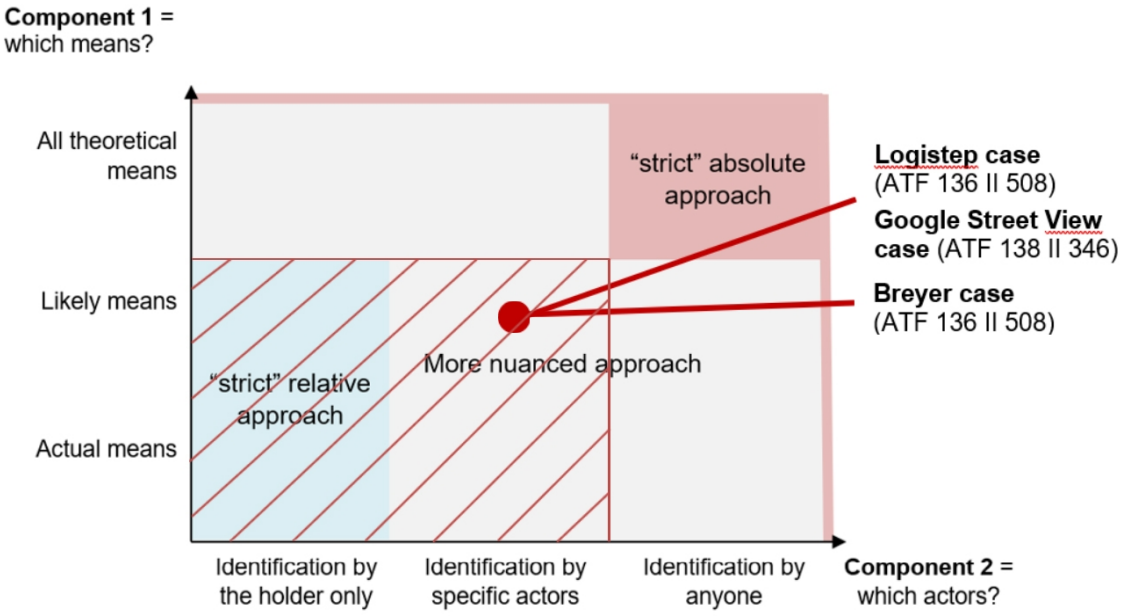

***Figure 3**: positioning CJUE Breyer on the «table of identifiability approaches» of Figure 1. The striped area shows all situations which would fall under the identifiability test of the CJUE.*

[68] Similarly to what applies under Swiss law, it is clear under the GDPR that the pseudonymized data remains personal data for any third party with **lawful means** to access the decryption keys.[107] Whether the pseudonymised data which arrives at another controller are still personal if the new controller does not have the legal or reasonably practical means to reverse the pseudonymization is controversial. Alongside other authors,[108] we consider that pseudonymized data can be anonymous data under the GDPR, but that this depends on the entire data situation (for the same reason as under Swiss law). This view is also shared by the UK data protection supervisory authority (ICO), which stated that pseudonymized data may become anonymous information in the hands of a third party under specific circumstances, which depend e.g. on the ability of the recipient to use other information to enable identification, the likelihood of identifiability, and the techniques and controls placed around the data once in the recipient's hands.[109] Nevertheless, this question remains heavily debated under the GDPR.[110]

107 Scheibner et al. (n 28), p. 6.

108 Groos/van Veen (n 44), p. 503; Scheibner et al. (n 28), p. 7.

109 ICO draft Anonymization, pseudonymisation and privacy-enhancing technologies guidance, Chapter 3: pseudonymisation (February 2022), p. 5.

110 For a more absolute approach, see Finck/Pallas (n 13), p. 19.

6. Considerations under the HRA

6.1. Concepts

[69] As specified above,[111] the HRA applies to the processing of health-related personal data in the context of medical research (Art. 2 para. 1 let. e HRA)[112]:

(a) The HRA defines «research» broadly as «methodological research aimed at obtaining generalizable knowledge».[113] The HRA applies not only to research activities carried out directly on persons, but also to research activities involving biological material of human origin or personal data (Art. 2 para. 1 let. e HRA).[114]

(b) Health-related personal data is defined by the HRA as «information concerning the health or disease of a specific or identifiable person, including genetic data». Except for the fact that the HRA specifically refers to health-related personal data, the notion of personal data in the HRA is similar to the one used in the FADP (Art. 3 let. b FADP) and covers both data relating to identified and identifiable persons.[115]

(c) The HRA **does not apply to research on anonymously collected or anonymised health-related data** (Art. 2 para. 2 let. c HRA). However, Art. 32 para. 3 HRA restricts the possibility to anonymize biological material and genetic data for research purposes. Anonymized health-related data is defined as health-related data which cannot (without disproportionate effort) be traced to a specific person (Art. 3 let. i HRA). Swiss law does not impose a specific test to determine if data is sufficiently anonymized. Art. 25 HRO requires that:

[...] all items which, when combined, would enable the data subject to be identified without disproportionate effort, must be irreversibly masked or deleted [...] In particular, the name, address, date of birth and unique identification numbers [...].

(d) The HRA applies to «coded» health-related personal data,[116] i.e. «data linked to a specific person via a code» (Art. 3 let. h HRA). Art. 26 HRO clarifies that:

1 *Biological material and health-related personal data are considered to be correctly coded in accordance with Article 32 paragraph 2 and Article 33 paragraph 2 HRA if, from the perspective of a person who lacks access to the key, they are to be characterized as anonymised.*

2 *The key must be stored separately from the material or data collection and in accordance with the principles of Article 5 paragraph 1, by a person to be designated in the application who is not involved in the research project.*

111 *Supra* chapter 2.

112 The HRA also applies to research activities carried out directly on humans.

113 Art. 3 lit. a HRO.

114 Federal Council Dispatch of 12 September 2007 concerning the constitutional article on research on human beings, FF 2007 6345, 6354.

115 Erard (n 10), p. 608; SHK HFG-van Spyk/Rudin/Sprecher/Poledna (n 10), Art. 3 N 43.

116 Valérie Junod/Bernice Elger, Données codées, non-codées ou anonymes, in: Jusletter 10 December 2018, N 6.

> (e) In other words, coding aims to replace certain identifying characteristics in such a way that only the person in possession of the code can re-identify the data.[117] It is therefore by definition reversible. This concept corresponds to the concept of pseudonymized data. Although the processes may be technically different, data pseudonymization and data coding are identical from a legal point of view.

[70] According to Junod/Elger and Erard, the contractual commitment of the data recipient not to attempt to re-identify the data has no influence on whether or not the data is anonymous or coded.[118] This view is in our opinion not entirely correct. As shown before, the risk of identification is heavily context-dependent and must be based on the entire data situation (data itself + its environment).[119] Furthermore, the assessment must consider not only the objective means available to a relevant person (objective assessment), but also her willingness to do so (subjective assessment).[120] In this respect, it can generally be assumed that, provided the data environment is controlled, the risk of re-identification in a medical research context remains relatively low (or at least lower than in other environments), considering the obligations imposed on researchers and consequences that an unauthorized re-identification operation would have on them (even if the risk that a researcher going astray and becoming an inside adversary can never be fully excluded).[121]

6.2. Discussion

[71] It is clear from the provisions cited above that, under the HRA, anonymization or pseudonymization does not necessarily have to be absolute.[122] Not all theoretical risks of re-identification will be relevant. The efforts required to identify must be assessed based on the test specified by the Swiss Federal Tribunal under the FADP.[123] In that respect, the HRA does not deviate from the regime under the FADP.

[72] In relation to the matter at stake, the main difference between the HRA and the FADP is that the HRA clearly specifies rules that also apply to coded data (and generally require the consent of the data subject)[124]. In particular, the HRA (Art. 32 and 33) institutes a complex regime of consent requirements for the secondary use of health-related personal data for research purpose, which can be summarized as follows:[125]

---

117 Meier (n 23), N 446 p. 206.

118 Erard (n 10), p. 609; Junod/Elger (n 116), N 11.

119 *Supra* chapter 3.3.

120 *Supra* chapter 3.4.

121 See Groos/van Veen (n 44).

122 The law already admits the effectiveness of the measure if re-identification requires «disproportionate efforts». Erard (n 10), p. 608.

123 *Supra* chapter 3.6.

124 Junod/Elger (n 116), N 6.

125 Junod/Elger (n 116), N 8; Erard (n 10), p. 607. The HRA is based on the principle of self-determination and privacy of the research participant, which is recognized as a fundamental pillar of research and is enshrined in the most important international texts on the subject (Erard (n 10), p. 614. E.g., Art. 5, 10 and 16 v) Convention on Human Rights and Biomedicine, SR 0.810.2; Arts. 9 and 32 World Medical Association Declaration of Helsinki (version 15 February 2017).

What type of data?	Non-Genetic Data	Genetic Data
Uncoded personal data	Generic consent	Specific consent
Coded data (pseudonymized)	Opt-out right (information, no objection)	Generic consent
Anonymized Data	HRA does not apply (research freedom)	Opt-out right (information, no objection)

[73] Article 32 and 33 HRA primarily serve to relax requirements of the HRA as regards the need to obtain consent. The HRA nevertheless gives the research subject a right of self-determination over her coded data. For example, the research subject must be able to limit the use of her coded data to a specific project.[126] There are other specific rules that apply to coded or non-coded personal data in the context of the HRA. As an example, according to Art. 43 HRA (and Art. 5 HRO), anyone who stores health-related personal data for research purposes (including in a coded form) must take appropriate technical and organizational measures to prevent unauthorized use thereof.

[74] Several authors have discussed the qualification of coded/pseudonymized data from the viewpoint of the researcher who does not have access to the key.

- Rudin considers that anyone who has (additional) knowledge on the basis of which the identity of the person concerned can be established (e.g. the treating and researching physician in the university hospital, who can easily establish the identity of the person suffering from a rare disease) can never invoke the fact that the data is anonymized in order to benefit from the corresponding privilege regarding justification for further use.[127]
- Junod/Elger leave this question unanswered[128] but consider that the researcher who receives a coded database for which he has no way of accessing the code (without having the means to identify the individuals) can only carry out his research project if the key holder contacts the subjects to inform them of their right to object (non-genetic data) or to obtain their consent (genetic data)–unless a «waiver» is granted by the ethics commission (Art. 34 HRA).[129]
- Erard considers that coded data must be regarded as personal data in the field of human research, even in relation to persons who do not have the key to re-identification. This solution is necessary in the field of research in view of the special regime imposed by the HRA and the particular rules imposed by this law for the processing of coded data (e.g. coded data may not in principle be used for any purpose other than research).[130] Erard considers that this justifies applying an «absolute approach» of the identifiability assessment in the context of medical research. The author argues that systematically considering that coded data is anonymous for the researchers who receive it (without the key), would result in the rules in the HRA (and the protection granted to the data subjects) being circumvented. In

126 Federal Council Dispatch HRA (n 94), 7337. See also: Lea Schläpfer, Clinical Data Sharing: Nutzen, Risiken und regulatorische Herausforderungen, in: recht 2016, pp. 136 ff, p. 141.

127 SHK HFG-Rudin (n 10), N 6 ad Art. 35.

128 Junod/Elger (n 116), N 16.5.

129 Junod/Elger (n 116), N 18.

130 On this particular issue, see: Erard (n 10), p. 606 ff.

> particular, once under the control of the recipients, the coded data would fall outside the scope of the HRA and/or the laws on data protection.[131]

[75] We concur with Erard that the position to be retained on this issue must not result in the law (and the protection it grants to the data subjects) being circumvented. Further, we note that because of the requirements of Art. 26 para. 2 HRO, data will only be validly coded if the researchers involved in the project do not have access to the re-identification key. If we consider that the data is fully anonymized from the point of view of those researchers, then the rules on the use of pseudonymized data would never apply.

[76] On the other hand, retaining a «strict» absolute approach could also result in unintended consequences (for the same reasons as explained under the FADP).[132]

[77] In our opinion, the same principles as outlined above under the FADP should also apply in the context of the HRA, i.e.:

- Identifiability must be assessed from the viewpoint of each data holder – i.e. the organization processing the data as controller.
- In the context of data sharing, the identifiability test must take into account the relevant data environment (from the position of the data holder).
- What is not permitted for the data holder should not become permissible only because it transfers pseudonymized or coded data to a third party which does not have the identification key.
- For a data recipient, the fact that the data provider retained a key (i.e. data is not anonymized but pseudonymized) is not by itself relevant. The assessment must be carried out based on the data recipient's point of view, taking into account its data environment.

Example 12

A researcher stores coded health-related personal data on a file-sharing SaaS application. Should the provider be subject to the FADP and the HRA? This is debatable. But what is clear is that the researcher cannot consider that this outsourcing is lawful on the sole basis that the information is coded and the service provider does not hold the key.

[78] A researcher who receives pseudonymized data will be directly subject to the rules of the HRA. If she can reasonably consider that the data is fully anonymized, she can process the data without further ado. If, on the contrary, she receives the data in the context of a research project and knows (or should know) that the data is only coded (because a third party retained the key),

131 Erard (n 10), p. 614–615.

132 *Supra*, chapters 3.4 and 3.5. Erard (n 10), p. 616 acknowledges this issue: «This raises the question of whether, in the research context, a ‹*heightened*› *relative approach should be considered, based not only on the effort calculation and interest in re-identification, but also on whether the researcher knows or has a reasonable suspicion that he or she is working with coded data. Where there is reasonable suspicion, which would be common in practice, such an approach would require the researcher to ascertain this from the source of the data and to take the necessary steps to comply with the statutory requirements for reuse.*»

then she will be directly subject to the HRA and may face criminal sanctions.[133] Conducting a research project without the required consent or the required authorization from the ethics committee constitutes a criminal offense only if the actor acted intentionally or negligently.[134]

[79] As is the case in the context of the FADP,[135] considering that the pseudonymized data may be handled as anonymized data from the viewpoint of the data recipient would entail additional obligations on the data provider.

Example 13

Art. 32 para. 3 HRA specifies that genetic data may only be anonymized if the data subject has been duly informed and did not object. If we consider that the communication of coded data without providing access to the key results in the data being anonymized from the point of view of the data recipient, then the data provider can only transfer the data in a coded form if it complies with the rules for anonymization.
This would mean that the data provider cannot transfer coded genetic data to third parties if the data subject objected to anonymization. To avoid this pitfall, the contract between the data provider and the data recipients should specify rules limiting the use of the data by the data recipient. In this manner, the data provider can ensure that the data will not be used by the data recipient as if it was anonymized data.

7. Conclusion

[80] The concept of personal data is as complex as it is important. Although many uncertainties remain, it is quite clear that the same piece of information may be personal data for one person, and an anonymous piece of data for another person. Anonymization (and pseudonymization from the viewpoint of the data recipient) is a heavily **context-dependent process** which requires consideration of the data situation as a total system, i.e. the data and its environment (people, other data and structure). *In case of data sharing, the identifiability test must* ***take into account the relevant data environment from the point of view of the data holder****.*[136] Identifiability must be assessed from the position of each data holder, but depends on the entire environment.

[81] The actual consequences of this situation – having data which is personal data for one party, but anonymous for another – must be analyzed on a case-by-case basis. In our opinion, the legal regime pertaining to the handling and transferring of pseudonymized (or coded) data should be based on the core principle that what is not permitted for the data controller should not become permissible only because it transfers pseudonymized or coded data to a third party which do not have the identification key. Accordingly, what a data controller cannot do, it cannot authorize (or permit) a third party to do.[137]

133 Junod/Elger (n 116), N 18.

134 Art. 62 para. 1 lit. a and b and para. 3 HRA; Art. 63 para. 1 lit. c HRA. On the criminal provisions of the HRA see SHK HFG-Gruberski (n 10), Vorbemerkungen Art. 62–64.

135 *Supra* chapter 4.

136 By «data holder», we mean the organization processing or controlling the data, and not individually any employee or department of that organization.

137 See Example 10 and Example 11. Although this principle may seem logical, applying a too narrow concept of identifiability would result in situations where the safeguards of data protection laws could easily be circumvented. For

[82] Determining that data is or is not, legally speaking, personal data for a data holder is important, but less than the actual consequences that this qualification entails. Contrary to what could be assumed, considering that the data is not personal data for the data recipient which receives pseudonymized data without the corresponding key may actually lead to additional obligations on the data provider. Indeed, in that case, in order for the data provider to ensure that it will be able to comply with its own obligations under the FADP, the data provider will have to impose appropriate obligations on the data recipient(s). To avoid situations in which the data provider cannot comply with its own obligations, the data provider will have to ensure that a contract appropriately protects its interests.[138] In other words, a contract often becomes more important, not less in such scenarios. Furthermore, two parties collaborating on a processing activity may be considered as joint controllers (both assuming obligations under data protection laws) even if only one of them has access to pseudonymized data, or even to no data at all.

Alexandre Jotterand, MLaw, attorney-at-law, CIPP/E, CIPM, id est avocats Sàrl (www.idest.pro).

instance, if the law provides for minimum security measures to protect pseudonymized research data stored internally, then the mere fact that the storage is outsourced to a third party should not result in a decrease of the level of security required.

138 See Example 06, Example 07, Example 08 and Example 09.