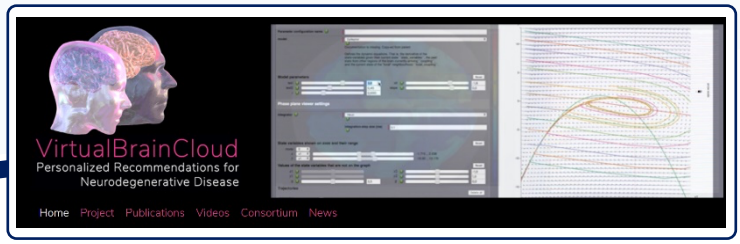
VirtualBrainCloud

Personalized Recommendations for Neurodegenerative Disease



www.VirtualBrainCloud-2020.eu

Public deliverable report

D2.5: Final analysis of relevant legal, ethical and regulatory framework, 2nd iteration, including a report on best practices and industry standards

Date	30 May 2023
Authors	Nikolaus Forgó, Katarzyna Barud, Michael Cepic, Emily Johnson, Mariana Rissetto (UNIVIE), Angela Bradshaw, Dianne Gove (AE), Petra Ritter (CHARITE), Pedro Omedas, Gerònimo Galìndez (Eodyne), Fraunhofer
Dissemination level	**public**
Website	www.VirtualBrainCloud-2020.eu

This project has received funding from the **European Union´s Horizon 2020** research and innovation programme under **grant agreement No 826421**

Table of content

Acronyms/Abbreviations

#	Acronym/ abbreviation	Description/Definition
	AD	Alzheimer's Disease
	DPIA	Data Protection Impact Assessment
	EC	European Commission
	EU	European Union
	GA	Grant Agreement
	GDPR	General Data Protection Regulation
	Individual and Patient	Both terms, in their singular and plural form, are in this report used interchangeably. However, Data Subject is defined in Art. 4 (1) of GDPR.
	NDD	Neurodegenerative Disease
	partners	Signatory parties/beneficiaries of the Grant Agreement No 826421
	Report	Deliverable 2.5
	TVB- Cloud	VirtualBrainCloud Project
	WP2	Work Package 2 TVB-Cloud
	WP29	Art. 29 Data Protection Working Party

1. Introduction

This Report (D2.5) provides the final analysis of relevant legal, ethical and regulatory framework, including a report on best practices and industry standards (MS 25)[1].

The Virtual Brain Cloud project (TVB-Cloud *or* project) envisions to enable personalized medicine that targets prevention, early diagnosis, disease progression prognosis, individual treatment plans and development of novel therapies for neurodegenerative diseases with focus on Alzheimer's and Parkinson's disease. To materialize this vision the project implements a European cloud-based platform that not only connects two critical streams of biomedical research, systems biology, and computational neuroscience, but that also connects clinics, researchers, patients, and students.[2] It comes to no surprise that this digital endeavor can only be achieved if personal data and special categories of such are

[1] This document also verifies Milestone 25.

[2] https://virtualbraincloud-2020.eu/tvb-cloud-the-project.html, accessed 15 September 2022.

processed. The use of such data within a cloud environment for health-related purposes as well as scientific research entails the application of a multi-layered legal framework.[3] This concerns foremost European data protection law and, therefore, the Regulation on the protection of natural persons with regard to the processing of personal data, well-known as GDPR[4]. Because of the current development status of TVB-Cloud, the following legal analysis and corresponding remarks deal with the project and its results more as a research tool than a health product. As well as complying with laws and regulations governing data privacy and health research, TVB-Cloud needs to address and mitigate ethical challenges linked to the development of brain simulations from multimodal health data using AI; and the sharing of data via the VRE, the cloud-based research environment created by the project. The following ethical analysis identifies key ethical challenges with relevance to TVB-Cloud, with recommendations on good practices for the project and its future iterations.

Deliverable D2.5 is the continuation and 2nd iteration of Deliverable D2.1[5]. It therefore builds up the aspects elaborated therein. The present work addressess the legal framework and related aspects in the first part and ethical framework and its related aspects in the latter.

2. Legal framework in TVB-Cloud Project

In this analysis, that constitutes a final analysis of relevant legal, ethical, and regulatory framework of TVB-Cloud as well as a report on best practices and industry standards, we follow a two-fold approach. On the one hand we want to retrospectively show certain legal issues that the project was confronted with during its development stage and the solutions achieved. This provides content for recommendations on best practice approaches from the legal perspective. This task was manly accieved by frequent exchanges with the developers of the project (especially via the legal help desk), by organizing and attending relevant conferences and by reaching out to external stakeholders in the field and bordering disciplines. On the other hand, we want to raise awareness for upcoming legislation that might become relevant for TVB-Cloud in the future. As a result of this two-fold approach, building upon the Deliverable D2.1, we display below the relevant aspects of the legal framework applicable to the TVB-Cloud.

2.1. Data Protection framework

[3] Michael Cepic, Mariana Risetto, Mapping the European Legal Framework on Security Requirements for Cloud Computing Infrastructures in the Healthcare Sector, EDPL vol. 4, 2020, p. 541.

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

[5] Grant Agreement number: 826421 — VirtualBrainCloud — H2020-SC1-DTH-2018-2020/H2020-SC1-DTH-2018, Annex I, Part B, p. 16.

The relevant data protection framework has already been identified in Deliverable D2.1. Deliverable D2.1. sets out that the GDPR, in principle, applies to the project and respective partners and that all pertinent provisions of said law must be complied with. In the course of the project, specific issues carose during the weekly technical or legal coordination meetings on the Virtual Research Environment (VRE, https://www.bihealth.org/de/translation/netzwerk/digitale-medizin/bihcharite-virtual-research-environment) that were dealt with via the legal help desk (Deliverable D2.3) in a memo format. These research memoscontain numerous research questions and have addressed the following issues:

- Brain Imaging and Data Protection
- Temporal and territorial scope of application of the GDPR
- Territorial scope GDPR
- Temporal scope GDPR
- Brexit update
- Storage limitation/Right to erasure GDPR
- Storage limitation GDPR
- Right to erasure GDPR
- Information duties GDPR

The detailed analysis of these issues and further considerations are detailed in the following subsections.

2.1.1. Brain Imaging and Data Protection

When determining material applicability of the GDPR, one of the first questions to ask is whether personal data is being processed. The GDPR only applies to personal data and not to non-personal data such as anonymised data or data unrelated to an individual (non personal data)[6]. Art. 4(1) of the GDPR defines personal data as meaning "any information relating to an identified or identifiable natural person", which it calls the 'data subject'. In determining identifiability, this action includes direct and indirect identifiability and examples are given in the GDPR, including a person's name, location data or, importantly for this discussion, "one or more factors specific to the physical [...] identity of that natural person".[7]

Moreover, the GDPR sets out additional rules for the processing of 'special categories of personal data'. Art. 9(1) GDPR prohibits the processing of special categories of personal data unless it is explicitly

[6] See also Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union.
[7] GDPR, Art. 4(1).

permitted under one of the ten exceptions listed in Art. 9(2) GDPR. A special category of personal data is that which is considered more sensitive because the processing of such data presents the possibility of a higher risk of harm to the data subject and their rights and freedoms. Special categories of personal data include genetic data, biometric data and data concerning health.

Despite all of the legal considerations, it is not always clear whether personal data is being processed. For example, the processing of brain images could provide details about someone's health, such as the presence of neurodegenerative conditions. However, the question of whether an individual can be identified from those images alone, and thus whether the GDPR applies, is a complex one. Authors Finck and Pallas have highlighted ambiguities in distinguishing personal data from non-personal data, stating that this binary view of the two data types "[in] reality operates on a different spectrum between data that is clearly personal data and that is clearly anonymous and anything in between".[8] Nevertheless, those processing personal data must make this distinction to remain lawful in their data processing.

The question then follows whether brain images are personal data within the meaning of Art. 4(1) GDPR. When determining whether brain images are personal data, there are several factors to consider. Firstly, is the brain image accompanied by other information. This information can include the patient name, medical ID or insurance number, their date of birth, contact details or even the presence of additional information on the image such as the individual's face. On their own, these data sets are personal data and when accompanying the brain image, they mean the brain image is associated with an identified or identifiable natural person, making the image personal data too. However, the question remains whether the brain image alone is personal data.

When determining the identifiability of a brain image alone, one can refer to the guidance given by the GDPR on assessing the possibility of re-identificantion for anonymous data. With the obvious identifiers removed the brain image alone could be rendered anonymous, therefore meaning the GDPR would not apply to subsequent data processing. However, there are a number of considerations that need to be taken in making this assessment and there is no definitive answer. Recital 26 of the GDPR requires the controller to make an assessment about the reasonable likelihood of identification, whether directly or indirectly, by them or another person. In assessing the likelihood of identification, the GDPR requires the controller to consider the following objective factors:

- Cost required for identification;
- Time required for identification;

[8] Michèle Finck and Frank Pallas, 'They Who Must Not Be Identified—Distinguishing Personal Data From Non-Personal Data Under the GDPR' (2020) 10(1) International Data Privacy Law 14.

- The available technology at the time of processing; and
- Future foreseeable technological developments.[9]

If it is ascertained that, taking into account all of these factors, the data subject cannot reasonably be identified, then it is not 'personal data' in accordance with the Art. 4(1) GDPR definition then the GDPR does not apply. However, the 'reasonable likelihood' of identification remains ambiguous and therefore burdensome in practice. As Edwards put it, "[w]hat constitutes personal data is one of the central causes of doubt in the current data protection regime".[10] Therefore, when processing brain images, particular attention needs to be given to the assessment of whether they constitute personal data, whether as a standalone image or with accompanying data. This is particularly crucial given the potential presence of special categories of personal data and the subsequent additional risks of harm to the data subject. In the field of complex health research, anonymization is virtually impossible if information relevant to the research purpose is to be kept in the data, so that identifiability for complex health data containing biometric information must always be assumed, especially for special categories of personal data.[11]

2.1.2. Temporal and territorial scope of application of the GDPR

As part of the TVB-C Indoc weekly technical meeting,[12] UNIVIE was asked two questions with relation to the extraterritorial and temporal scope of application of the GDPR. The questions were as follows:

a. How does the GDPR apply to data that was collected in the US from US citizens that will now be used in the VBC infrastructure? (Territorial scope)

b. Personal data that was collected before the adoption of the GDPR (for example in 2008) is planned on being used in the VBC project. What data protection rules apply to these datasets?

2.1.2.1. Territorial scope

Art. 3 (1) GDPR sets out that the GDPR "*applies to the processing of personal data in the context of the activities of an establishment of a controller or a processor in the Union, regardless of whether the processing takes place in the Union or not.*"

[9] GDPR, Recital 26.
[10] Lilian Edwards, 'Data Protection I: Enter the GDPR' in Lilian Edwards (ed), Law, Policy and the Internet (Hart 2018) 84.
[11] Mary Donelly and Maeve McDonagh, 'Health Research, Consent and the GDPR Exemption' (2019) European Journal of Health Law, 26, 97-119 (100).
[12] Memo dated 19.10.2020.

Given the wording of Art. 3(1) GDPR and the fact that the VBC both is located within the European Union (EU) and the data processing will be carried out within the EU, the GDPR will apply to the processing of all personal data processed for the VBC. This includes personal data collected in the US from US citizens by virtue of the principle of establishment set out in Art. 3(2)(a) GDPR. This means, that because the VBC, and the institutions operating this project, are situated within the EU, the GDPR applies to them.[13] Furthermore, with regard to the previous instrument that regulated data protection in the EU prior to the adoption of the GDPR, Directive 95/46/EC, the European Court of Justice[14] ruled that the territorial scope of the Directive must be interpreted extensively in order to guarantee *effective and complete protection of the fundamental rights and freedoms of natural persons, and in particular their right to privacy.*[15] To emphasise the strictness of the principle of establishment, the GDPR also applies to institutions that outsource data processing to other countries in the world when processing the personal data of EU data subjects. The means and purposes of the processing are irrelevant to the territorial applicability of the GDPR; when the controller or processor is located in the EU, the GDPR is applicable.[16] Consequently, the processing of the personal data of US citizens by the VBC in the EU comes under the scope of the GDPR.

2.1.2.2. Temporal scope

Art. 99 (1) & (2) GDPR state:

This Regulation shall enter into force on the twentieth day following that of its publication in the Official Journal of the European Union. It shall apply from 25 May 2018.

First, it needs to be noted that any past, present or prospective processing of personal data, that falls within the temporal scope of the GDPR, so 25 May 2018 and onwards, needs to be GDPR compliant. Thus, when a data set of personal data which was collected prior to the implementation of the GDPR is used at the present time, or any time in the future, controllers and processors are required to comply with the GDPR.

Recital 171 sentence 3 states: *Where processing is based on consent pursuant to Directive 95/46/EC, it is not necessary for the data subject to give his or her consent again if the manner in which the consent*

[13] See also EDPB Guidelines 3/2018 on the territorial scope of the GDPR (Art. 3) (12.11.2019), 5 ff: https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines_3_2018_territorial_scope_after_public_consultation_en_1.pdf, accessed on 16 October 2020.
[14] Case C-131/12 *Google Spain SL and Google Inc. v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González* [2015] (ECLI:EU:C:2014:317).
[15] *Urecker*, Extraterritorialer Anwendungsbereich der DS-GVO - Erläuterungen zu den neuen Regelungen und Ausblick auf internationale Entwicklungen, ZD 2019, 67 (67 ff).
[16] *Ernst* in *Paal/Pauly* (Eds) DS-GVO BDSG (2nd edition 2018) margin number 11 f.

has been given is in line with the conditions of this Regulation, so as to allow the controller to continue such processing after the date of application of this Regulation.

Recital 171 reflects the requirement of informed consent set out in the Directive whereby Art 2(h) on the definition of 'the data subject's consent' requires that consent must be freely given, specific and informed. Art. 7 and Recital 32 GDPR, like the Directive require that consent be freely given, specific and informed. The only additional requirements of the GDPR not explicitly required by the Directive are that consent should be a clear affirmative act and unambiguous intention of the data subject's consent.[17] However, the Directive does require 'explicit consent'.[18] While the two pieces of legislation feature slightly differing wordings, the substance of the requirements is very similar, and when read in light with Recital 171 (above) of the GDPR, we can conclude that consent given **before** the GDPR entered into force, which is in line with the Directive, can be compatible with the requirements for consent set out in the GDPR.[19]

It must, however, be noted that informed consent given under the Directive only remains valid, according to one decision of the *Düsseldorfer Kreis*, if it was given freely[20] and by persons above the age of 16[21].[22] According to the last sentence of Art 8(1) GDPR "*Member States may provide by law for a lower age for those purposes provided that such lower age is not below 13 years."* In case a Member State made use of this provision (opening clause), you should check, whether consent given under the Directive also fulfils the national age limit, that is applicable now under the national transposing law of Art. 8(1) GDPR.

If consent was given in accordance with a national law that transposed Directive 95/46/EC, which should be the case with any consent given in the last two decades, the consent should be in line with the same requirements set out in the GDPR. Directive 95/46/EC, Art. 7(a) provided the equivalent basis for consent: "*the data subject has unambiguously given his consent*", whereas the GDPR, Art. 6(1)(a) says "*the data subject has given consent to the processing of his or her personal data for one or more specific purposes;*". Even with the omission of the purpose principle in Art. 7 of the Directive, given that the Directive also sets down the purpose limitation in Art. 6(1)(b), we see no problem for the use of that data. Notably, any secondary use as well as all other processing operations that take place presently

[17] GDPR, Recital 32.
[18] Directive 95/46/EC, Recital 33.
[19] Tinnefeld/Conrad, ‚Die selbstbestimmte Einwilligung im europäischen Recht – Voraussetzungen und Probleme', ZD 2018, 391 (395).
[20] GDPR, Art. 7(4) in connection with Recital 43.
[21] GDPR, Art. 8(1) in connection with Recital 38.
[22] Beschluss der Aufsichtsbehörden für den Datenschutz im nicht-öffentlichen Bereich (Düsseldorfer Kreis am 13./14. September 2016); see also: <https://datenschutz.hessen.de/infothek/duesseldorfer-kreis> accessed 22 September 2022.

and/or prospectively, must comply with all provisions of the GDPR. Secondary use, or further processing means and personal data which were collected for one purpose and are now being used for a different purpose. Any personal data being processed for secondary purposes, must also be processed in line with the GDPR. Thus, the personal data previously collected under the Directive and now being used for different purposes, must be GDPR compliant and this applies to the consent requirements mentioned above. If the consent obtained previously is compliant with the Directive, then it will likely be compliant with the GDPR as set out in Recital 171 GDPR.

On the other hand we also want to share, that some commentators[23] suggest that because the GDPR allows Member States to *individually maintain or introduce further conditions, including limitations, with regard to the processing of genetic data, biometric data or data concerning health,*[24] that this could be indicative for a need to actualise old consent forms, when consent was given to the processing of such categories of data and a Member state has made use of that opening clause. Furthermore, some authors suggest that because of the non-binding character of the Recitals, increased caution should be exercised, when the compatibility of old consent forms (in accordance with the old Directive) with the GDPR is assessed.[25] Lastly, there is an open debate on whether data subjects that consented according to the old Directive were able to compatibly do so with the (now applicable) GDPR when it comes to the information duties.[26] In other words, whether consent under the old legal regime can be interpreted as having covered the same information as is now required by the GDPR.

2.1.2.3. Conclusion

We come to the conclusion that any institution that is situated in the EU needs to comply with the GDPR, when they process personal data, regardless of the origin of those data (principle of establishment). Furthermore, we suggest that when processing personal data that was collected before the GDPR entered into force, a case by case assessment of the consent forms used needs to be made to examine whether the consent obtained is in line with the GDPR requirements. Especially since there are diverging views on the compatibility of consent given under the scope of the Directive with the requirements of the GDPR we suggest that a case by case analysis is performed and contact with UNIVIE and your DPO is established.

2.1.2.4. Brexit update

[23] E.g. Pauly in Paal/Pauly (eds) *DS-GVO BDSG* (2nd edition 2018) margin number 6.
[24] GDPR, Art. 9(4).
[25] Heckmann/Paschke 'Artikel 7 Bedingungen für die Einwilligung' in Ehmann/Selmayr (eds), *Datenschutz-Grundverordnung* (2nd edition 2018), margin number 101.
[26] Specht/Mantz in *Handbuch Europäisches und deutsches Datenschutzrecht* (1st edition 2019), Fortgeltung von Alteinwilligungen, margin number 45.

As mentioned in the previous iterations of this deliverable, Art. 3 GDPR confirms that the territorial scope of GDPR covers the processing of personal data in the context of activities "of an establishment of a controller or processor in the Union, regardless of whether the processing takes place in the Union or not".[27] As such, the GDPR may also be applicable to third-countries. Chapter 5 of the GDPR stipulates the additional legal considerations applicable when personal data is transferred to third countries.

One of the additional legal measures for lawful third-country data transfers is on the basis of an adequacy decision. Adequacy decisions are decisions made by the European Commission based on Art. 45 GDPR, in which the data protection framework of a third country is assessed, then the European Commission has decided that a given third country has an adequate level of data protection law at least equivalent to EU data protection law. Consequently, adequacy decisions allow for the flow of personal data from the EU (Iceland, Liechtenstein and Norway) to a third country without the need for additional safeguards. To date, the European Commission has recognised thirteen countries as having an adequate level of data protection law.[28]

As of 28 June 2021, the UK received an adequacy decision from the European Commission. As highlighted by the EDPB, the UK is under the jurisdiction of international human rights law, including the European Convention on Human Rights as well as Convention 108 and it has signed Convention 108+.[29] The integration of EU data protection law within the UK's legal system as well as international legal obligations contributed to the granting of the adequacy decision of the GDPR and LED. Věra Jourová, the Vice-President for Values and Transparency, stated that although the UK has left the EU, "today its legal regime of protection of personal data is as it was".[30] Consequently, post-BREXIT data transfers from the EU to the UK are permitted and any such transfer "shall not require any specific authorisation".[31]

2.1.3.Storage limitation/Right to erasure

Another aspect brought to our attention and subject of analyisis was the legal requirements for the storage and erasure of collected personal data for research purposes.[32] For this, we need to highlight that while storage and erasure are connected, they must also be differentiated, as both are regulated

[27] GDPR, Art. 3(1).
[28] European Commission, *Adequacy Decisions* <https://ec.europa.eu/info/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions_en> accessed 4 April 2022.
[29] Protocol amending the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data ("Convention 108+"), 18 May 2018.
[30] European Commission, *Data Protection: Commission Adopts Adequacy Decisions for the UK* (European Commission, 28 June 2021) <https://ec.europa.eu/commission/presscorner/detail/en/ip_21_3183> accessed 17 November 2021.
[31] GDPR, Art. 45(1).
[32] Memo dated 23 July 2020.

separately. For this reason, we first outline the principle of storage limitation and continue with the right to erasure.

2.1.3.1. Storage limitation

Art. 5 (1)(e) GDPR sets out the 'storage limitation' principle requiring that personal data shall be

> *"kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed; personal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) subject to implementation of the appropriate technical and organisational measures required by this Regulation in order to safeguard the rights and freedoms of the data subject ('storage limitation')."*

The storage limitation is an essential requirement of the data protection in the EU. Recital 39 of the GDPR further specifies: *"[…] This requires, in particular, ensuring that the period for which the personal data are stored is limited to a strict minimum. Personal data should be processed only if the purpose of the processing could not reasonably be fulfilled by other means. In order to ensure that the personal data are not kept longer than necessary, time limits should be established by the controller for erasure or for a periodic review […]."*

In assessing adherence to the storage limitation principle, controllers should ask the following questions:[33]

- Does any storage of personal data take place in the context of the specific processing operation?
- Should this be the case, what is the storage period that is necessary for the purpose? Can it be justified in relation to the purpose of the processing?
- Does the storage period vary between different data items? In such cases, the controllers should try to define (and justify) specific storage periods for different data items.
- Are data erased after the end of their defined storage period? If this is not the case, the controller should explain the purpose for which the data are further processed, how storage is performed (location, recipient) and the planned retention period.

[33] Cf. ENISA, *Recommendations on shaping technology according to GDPR provisions* (2018) <https://op.europa.eu/en/publication-detail/-/publication/a8e7a463-29c5-11e9-8d04-01aa75ed71a1/language-en> accessed on 13 July 2020.

The determination of the time limits and criteria requires a case-by-case consideration, in which the necessity of the retention of data is assessed on the basis of the processing purposes.[34] This includes potential data protection related obligations to have data stored (such as the compliance with the principle of accountability) but also other legal obligations for an extended storage period, for example due to tax law provisions or health law provisions. The data may also be stored longer[35] if they are processed for archiving, scientific and historical research and statistical purposes in the public interest.[36] In this regard you will need to consider the necessary timeframe of storage for the specific purposes of the research conducted. As a rule of thumb, you will need to cease storing personal data, if they are no longer (absolutely) necessary for the research you are conducting. The data controller will also need to consider, whether anonymised data would be sufficient for the intended research, in which case the storage of non-anonymised (personal) data would need to cease.[37] In the end it is up to the data controller to determine what storage period is necessary and proportionate to perform the specific research.

2.1.3.2. Right to erasure

Corresponding to Art. 5 (1)(e) GDPR is Art. 17 GDPR that regulates the right of erasure. It states that:

> *"1. The data subject shall have the right to obtain from the controller the erasure of personal data concerning him or her without undue delay and the controller shall have the obligation to erase personal data without undue delay where one of the following grounds applies:*
>
> *(a) the personal data are no longer necessary in relation to the purposes for which they were collected or otherwise processed;*
>
> *(b) the data subject withdraws consent on which the processing is based according to point (a) of Article 6(1), or point (a) of Article 9(2), and where there is no other legal ground for the processing;*

[34] Walter Hötzendorfer et al. 'Art 5 DSGVO. Grundsätze für die Verarbeitung personenbezogener Daten', in Rainer Knyrim (edt.), *DatKomm* (2018) margin number 50.
[35] In theory data can be stored indefinitely, if they remain essential to a specific continuous research activity, see also <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/principles/storage-limitation> accessed on 15 July 2020.
[36] Walter Hötzendorfer et al. 'Art 5 DSGVO. Grundsätze für die Verarbeitung personenbezogener Daten', in Rainer Knyrim (edt.), *DatKomm* (2018) margin number 52.
[37] Peter Schantz ‚Art 5 DSGVO', in Wolff/Brink (eds), *BeckOK Datenschutzrecht* (2020) margin number 34.

(c) the data subject objects to the processing pursuant to Article 21(1) and there are no overriding legitimate grounds for the processing, or the data subject objects to the processing pursuant to Article 21(2);

(d) the personal data have been unlawfully processed;

(e) the personal data have to be erased for compliance with a legal obligation in Union or Member State law to which the controller is subject;

(f) the personal data have been collected in relation to the offer of information society services referred to in Article 8(1).

[…]

3. Paragraphs 1 and 2 shall not apply to the extent that processing is necessary:

(a) for exercising the right of freedom of expression and information;

(b) for compliance with a legal obligation which requires processing by Union or Member State law to which the controller is subject or for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller;

(c) for reasons of public interest in the area of public health in accordance with points (h) and (i) of Article 9(2) as well as Article 9(3);

***(d) for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) in so far as the right referred to in paragraph 1 is likely to render impossible or seriously impair the achievement of the objectives of that processing**; or*

(e) for the establishment, exercise or defence of legal claims"

There are several instances in which a data controller would need to either erase data on their own or at a data subject's request. Relevant for research, however, is the exception of Art. 17(3)(d) GDPR. As outlined by Quinn & Quinn (2018),[38] erasing personal data that is collected for research purposes right after their initial processing would prohibit any further research. Hence, the existence of the research exception. The exception is subject to the condition that erasure is likely to make it impossible or

[38] Paul Quinn & Liam Quinn, 'Big genetic data and its big data protection challenges' (2018) Computer Law & Security Review 34 1000-1018.

seriously prejudicial to the achievement of the purposes of the data processing in question. This may be the case if the results of the data processing operations would be significantly distorted by the deletion of some data sets, in particular in the case of small samples. Due to the reference to Art. 89(1), the applicability of the exception also requires that the conditions of Art. 89(1) are fulfilled, i.e. that the processing operations for the purposes mentioned are subject to adequate safeguards for the rights and freedoms of the data subjects.[39] Furthermore, if partners rely on the exception of Art. 17(3)(d) GDPR, they will need to check (before), whether the non-erasure of personal data is necessary and proportionate.[40]

Although the GDPR sets accountability as an overarching outcome of adhering to the Art. 5(1) principles, and it has recognised the problem of the obligation to delete and retain documents (and the personal data therein) in Art. 17(3)(b) itself, the GDPR does not set any deadlines for the retention of certain documents and therefore does not provide a specific period after which data need to be deleted.[41]

2.1.3.3. Information duties:

In principle there is also the duty of the controller to inform the data subjects on the storage period of their personal data (Art. 13(2)(a) & Art. 14(2)(a) GDPR). However, we can read from Recital 62 of the GDPR that "*it is not necessary to impose the obligation to provide information […] where the provision of information to the data subject proves to be impossible or would involve a disproportionate effort. The latter could in particular be the case where processing is carried out for […] scientific research purposes.* ***In that regard, the number of data subjects, the age of the data and any appropriate safeguards should be taken into consideration.***"

2.1.3.4. Conclusion:

For a research project like the TVB-C, the requirement of storage limitation will need to be assessed individually on a case by case basis (e.g each dataset individually), keeping in mind which data needs to be kept in order to be able to continue the intended research. Therefore, partners will need to consider the purpose for which the data were collected, the length of the research and its future purpose.[42] They will also need to consider if the data are essential for the ongoing research activity and that safeguards

[39] Viktoria Haidinger 'Art 17 DSGVO. Recht auf Löschung', in Rainer Knyrim (edt.) *DatKomm* (2018) margin number 73.
[40] Christopher F. Mondschein & Cosimo Monda, 'The EU's General Data Protection Regulation (GDPR) in a Research Context', in Pieter Kubben et al. (eds), *Fundamentals of Clinical Data Science* (2019) 55 (65-68).
[41] Sven Hunzinger, 'Löschkonzepte nach der DSGVO am Beispiel von ERP-Systemen' (2018) Computer und Recht 6 357-366.
[42] European Data Protection Board, *Guidelines 03/2020 on the processing of data concerning health for the purpose of scientific research in the context of the COVID-19 outbreak* (2020) < https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines_202003_healthdatascientificresearchcovid19_en.pdf> accessed 16 July 2020.

as outlined in Art. 89 of the GDPR are implemented. Lastly, they should also check, whether national law on research activities governs a specific storage length.[43] As an example, in Austria the *Forschungsorganisationsgesetz*[44] ("Research organisation act") mandates how long raw data can be stored.

In the event that a data subject exercises their right to erasure, partners will need to inform the data subject that they refuse to comply with the request on the grounds of the research exception. They will also need to inform them of their right to make a complaint to a supervisory authority and their ability to seek to enforce this right through a judicial remedy.[45]

We were able to establish that specific research purposes may provide a legal reason to keep data stored for an extended storage period. However, this must be able to be clearly demonstrated. When relying on this partners will need to keep in mind that the safeguards required by Art. 89 GDPR will need to be implemented in addition to all of the other principles of data processing as set out in Art. 5 GDPR. For further information to the Art. 5 principles please see Deliverable D2.1.

2.1.4. Regulation 2018/1807 (Non-personal data regulation) and concept of anonymization/pseudonymisation from a legal perspective

Following our TVB-C VRE legal meeting on February 11, 2021 we were confronted with the question of whether animal data can be processed in the VRE and to outline the legal framework on non-personal data.[46] In the pursuit to generally assess this question we will also cover aspects of anonymization and pseudonymisation.

2.1.4.1. General assessment

To ensure a free flow of data other than personal data (non-personal data) within the European Union (EU), Regulation (EU) 2018/1807 (Non-personal data regulation) was passed.

According to Art. 1 of Regulation 2018/1807, this Regulation aims to ensure the free flow of data other than personal data in the EU, and to do this it lays down the rules relating to data localization requirements, the availability of data to competent authorities and the porting of data for professional

[43] Next to the proportionality and necessity principles Recital 156 of the GDPR stresses, that "[*T]he processing of personal data for scientific purposes should also comply with other relevant legislation such as on clinical trials*."
[44] See specifically § 2f (3) Bundesgesetz über allgemeine Angelegenheiten gemäß Art. 89 DSGVO und die Forschungsorganisation in its effective version BGBl (Federal law gazette) I 31/2018.
[45] Available at <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/individual-rights/right-to-erasure/> accessed 15 July 2020.
[46] Memo dated 12 February 2021.

users. As set out in Art. 2, the Scope of this Regulation applies to the processing of electronic data other than personal data in the Union, which is:

> *(a) provided as a service to users residing or having an establishment in the Union, regardless of whether the service provider is established or not in the Union; or*
>
> *(b) carried out by a natural or legal person residing or having an establishment in the Union for its own needs.*
>
> *2. In the case of a data set composed of both personal and non-personal data, this Regulation applies to the non-personal data part of the data set. Where personal and non-personal data in a data set are inextricably linked, this Regulation shall not prejudice the application of Regulation (EU) 2016/679. 3. This Regulation does not apply to an activity which falls outside the scope of Union law.*

The definition of non-personal data follows an *e contrario* approach meaning that all data not being personal data (as defined in Art. 4(1) of the GDPR) are non-personal data.[47] Therefore, animal data are considered non-personal data and are regulated under the scope of Regulation 2018/1807.

Non-personal data can be divided into two categories:

> *1. data which originally did not relate to an identified or identifiable natural person, such as data on weather conditions generated by sensors installed on wind turbines or data on maintenance needs for industrial machines, or animal data; and*
>
> *2. data which were initially personal data, but were later made anonymous. The 'anonymisation' of personal means that the data cannot be attributed to an identified or identifiable person either directly or indirectly. It must not be possible to identify an individual even with the use of additional data. As such, anonymised data are non-personal data.[48]*

Examples for non-personal data are:[49]

> · *Data which are aggregated to the extent that individual events (such as a person's individual trips abroad or travel patterns which could constitute personal data) are no longer identifiable, can be qualified as anonymous data.*

[47] European Commission, *Guidance on the Regulation on a framework for the free flow of non-personal data in the European Union*, COM (2019) 250 final.
[48] Ibid.
[49] Ibid., 6 f.

· *Anonymous data are used or instance in statistics or in sales reports (for example to assess the popularity of a product and its features). High-frequency trading data in the finance sector, or data on precision farming which help to monitor and optimise the use of pesticides, nutrients and water.*

Note: In contrast to anonymised data (not personal data), pseudonymised data is personal data. The process of pseudonymisation is:

> *the processing of personal data in such a way that it is not possible to attribute them to a specific person without the use of additional information. This additional information is kept separately and is secured through organisational or technical measures (e.g. encryption). Nonetheless, data which have been pseudonymised are still considered information about an identifiable person if they can be attributed to this person by using additional information. Such data constitute personal data in accordance with the General Data Protection Regulation.*[50]

> *Retraceably pseudonymised data may be considered as information on individuals which are indirectly identifiable. Indeed, using a pseudonym means that it is possible to backtrack to the individual, so that the individual's identity can be discovered, but then only under predefined circumstances. In that case, although data protection rules apply, the risks at stake for the individuals with regard to the processing of such indirectly identifiable information will most often be low, so that the application of these rules will justifiably be more flexible than if information on directly identifiable individuals were processed.*[51]

To regard anonymised and pseudonimsed data to be of equivalent or even similar in nature is a common misconception in data protection.

> *Pseudonymised data cannot be equated to anonymised information as they continue to allow an individual data subject to be singled out and linkable across different data sets. Pseudonymity is likely to allow for identifiability, and therefore stays inside the scope of the legal regime of data protection. This is especially relevant in the context of scientific, statistical or historical research.*[52]

Additionally, the purposes of each process are different: pseudonymisation is used as a safeguard in the processing of personal data, anonymisation on the other hand transforms personal data to non-personal data.

[50] Ibid.
[51] Art. 29 Working Party, *Opinion 4/2007 on the concept of personal data*, 01248/07/EN WP 136, 18.
[52] Art. 29 Working Party, *Opinion 05/2014 on Anonymisation Techniques*, 0829/14/EN WP 216, 10.

To provide an example of the possible pitfalls of the misconceptions surrounding the technique of pseudonymisation can be found in the *"AOL (America On Line) incident"* where:

> *In 2006, a database containing twenty million search keywords for over 650,000 users over a 3-month period was publically released, with the only privacy preserving measure consisting in replacing AOL user ID by a numerical attribute. This led to the public identification and location of some of them. Pseudonymised search engine query strings, especially if coupled with other attributes, such as IP addresses or other client configuration parameters, possess a very high power of identification.*[53]

In determining whether the anonymisation process has been successful, and subsequently whether an individual is either directly or indirectly identifiable, controllers and processors must objectively consider means reasonably likely to be used identify the individual. Recital 26 of the GDPR states that:

> *To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments.*[54]

2.1.4.2. Conclusion

For VBC and the VRE, Regulation (EU) 2018/1807 (Non-personal data regulation) will be applicable as it provides a digital service (partially) with non-personal data and Charité or the infrastructure provider is residing in the EU. The Regulation applies to non-personal data and therefore also to a data set, that contains mixed data of personal and non-personal nature. To outline an example for such a mixed data set, it can be state that:

"[H]ealth data can be part of a mixed dataset. Examples include electronic health records, clinical trials or sets of data collected by various mobile health and wellbeing applications (such as applications for measuring our health status, for reminding us to take our medication or for tracking our fitness progress). The exact division between personal and non-personal data in these datasets is becoming increasingly blurred with technological developments. Consequently, their processing must comply with the General Data Protection Regulation, in particular (given that health data is a special category of data according to the Regulation) with Article9, which lays out a general prohibition on the processing of special categories of data and exceptions from this prohibition. The data in mixed datasets containing health data can be a valuable source of information, e.g. for further medical research, for measuring the side

[53] Ibid, 11.
[54] GDPR, Recital 26.

effects of a prescribed medicine, for disease statistical purposes or for developing new healthcare services or treatments. However, the General Data Protection Regulation must be complied with when carrying out the initial processing operations and when carrying out further data processing operations. Therefore, any such processing of health data must have a valid legal basis and an appropriate justification, be secure and provide for sufficient safeguards."[55]

The GDPR must be complied with in regard to the processing of the personal data. The main purpose of the Regulation EU 2018/1807 is the free flow of data. This Regulation acknowledges existing legislation, namely the GDPR and the higher level of protection accorded to personal data.

The VRE seems to be compatible with Regulation (EU) 2018/1807 (Non-personal data regulation).

If it is confirmed that no personal data will be processed in the VRE, then Regulation EU 2018/1807 applies and the processing of animal data or non-personal test data within the VRE may go ahead. Nevertheless, we recommend seeking confirmation from your organisation's DPO and/or legal team.

2.1.5. GDPR roles in the context of the relationship employer/employee

The following questions were raised with regard to the GDPR roles in the context of the relationship employer/employee:[56]

a. When academic employees leave an institution. How is data controllership handled? What is the role of the institution within GDPR?

b. Is the scientist the data controller – not the institution? In what cases is the institution the data controller?

c. Will /should in future increasingly the institutions take the role of data controllers – rather individual scientists?

2.1.5.1. GDPR roles / relationship employer/employee

The roles determined in the GDPR are regulated under Chapter 4 GDPR. The definition of controller and processor is as follows:

[55] Guidance on the Regulation on a framework for the free flow of non-personal data in the European Union, COM (2019) 250 final, 10.
[56] Memo dated 3.8.2021.

- **Controller**: means the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data (Art. 4(7) GDPR)
- **Processor**: means a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller (Art. 4(8))

The underlined parts highlight the aspect that makes these roles distinct. A detailed explanation of these roles and interaction was made available in the TVB webinar, delivered on 14 May 2020,[57] which is here use as point of reference.

In order to understand the relationship between employer/employee in the context of GDPR and its roles, two cases must be detailed:

2.1.5.2. Employees acting under the direct authority of controller

The Guidelines 07/2020 on the concepts of controller and processor in the GDPR issued by the EDPB ('the Guidelines 07/2020') state that "whereas the terms […] 'controller' and 'processor' are defined in the Regulation, the concept of 'persons who, under the direct authority of the controller or processor, are authorised to process personal data' is not. It is, however, generally understood as referring to persons that belong to the legal entity of the controller or processor (an employee or a role highly comparable to that of employees, e.g. interim staff provided via a temporary employment agency) but only insofar as they are authorized to process personal data."[58]

Therefore, three facts must be stated:

- Employees are generally understood to be "persons who, under the direct authority of the controller or processor, are authorised to process personal data";
- GDPR does not define such a category, therefore employees who,
    1. are under the direct authority of the controller or processor, and
    2. are authorised to process personal data

are neither a controller or processor, but understood they belong to legal entity.

[57] TVB-Cloud Webinar on data sharing in health research – EU research projects- TVB-Cloud Project, 14 May 2020. Available at <https://www.youtube.com/watch?v=X4bX3EfoSsU&t=224s> accessed 22 September 2022.
[58] European Data Protection Board, *Guidelines 07/2020 on the concepts of controller and processor in the GDPR* Version 2.0, adopted on 07 July 2021, <https://edpb.europa.eu/system/files/2021-07/eppb_guidelines_202007_controllerprocessor_final_en.pdf> accessed 29 November 2022.

To reinforce this statement, the EU Commission explains under the question 'what is a data controller or a data processor?' that

> "[…], if your company/organisation decides 'why' and 'how' the personal data should be processed it is the data controller. Employees processing personal data within your organisation do so to fulfil your tasks as data controller".[59]

2.1.5.3. Employees NOT acting under the direct authority of controller

The Guidelines 07/2020 state that 'an employee who obtains access to data that he or she is not authorised to access and for other purposes than that of the employer' does not fall within the category [of *persons who, under the direct authority of the controller or processor, are authorized to process personal data]'*.

On the other hand, 'this employee should be considered as a third party vis-à-vis the processing undertaken by the employer. Insofar as the employee processes personal data for his or her own purposes, distinct from those of his or her employer, he or she will then be considered a controller and take on all the resulting consequences and liabilities in terms of personal data processing'[60]

2.1.5.4. Conclusion

In conclusion, and to shed light on the questions posed,

1. In the case described in 2.1.5.2. the employee acts under the authority of the employer, being a person who is authorized to process personal data. Therefore, prima facie, there is no controllership relationship to be dealt with in the case of an employee leaving an institution. The dissolution of such relationship is regulated under the law/rules applicable to such employment relationship. In the case described in 2.1.5.3., a case-by-case analysis should be performed.

2. The role of the institution in the context of the GDPR should be subject to a case-by-case analysis of the definition of Art. 4 (7) and Art. 4 (8) GDPR.

[59] European Commission, 'What is a data controller or a data processor?' https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/obligations/controller-processor/what-data-controller-or-data-processor_en accessed 29. November 2022.
[60] Ibid, 2.

3. If there is an employment relationship between a scientific researcher and the research institution, the role of the first will be determined depending whether it acts under the direct authority of the controller or not (see case 2.1.5.2. and 2.1.5.3.).

2.1.6. Data sharing

Data sharing within TVB-C project is goverened in accordance with the GDPR. In this regard, several steps were taken regarding data sharing and its legal pursue within the Project.

1) UNIVIE provided the partners with a use case scenario Use Case Scenario for personal data sharing as an annex to D3.4 ('Interim source-space multiresolution MEG time series in BIDS share. Initial MEG and SEEG brain dynamic measures at the disposal of other WPs') (See Annex II), where aspects of data sharing were detailed, an inquiry for the data flow was dealt with and recomendataions were drawn. (December 2019)
2) TVB-Cloud Webinar on data sharing in health research – EU research projects- TVB-Cloud Project, 14 May 2020.[61]
3) Data sharing agreements negotaitions. Where applicable, data sharing agreements where signed. These were developed in particular for specific data sharing operations between partners and in general as part of the VRE platform, that succesfully underwent an external legal review.[62]

2.2. Cloud computing legal framework

This section focuses on the European legal framework relevant to cloud computing infrastructure deployed in the healthcare sector, such as the TVB-Cloud. During the TVB-Cloud project, three main areas important for security requirements for the cloud computing have been identified and further explored by the project: data protection, healthcare, and cybersecurity. These frameworks have been identified in Deliverable 2.1 and further discussed in the publication written by Michael Cepic and Mariana Rissetto (UNIVIE)[63] and are further explained.

[61] TVB Webinar <https://www.youtube.com/watch?v=X4bX3EfoSsU&t=224s> accessed 22 September 2022.
[62] BIH/Charité Virtual Research Environment <https://www.bihealth.org/de/translation/netzwerk/digitale-medizin/bihcharite-virtual-research-environment?tx_news_pi1%5Baction%5D=detail&tx_news_pi1%5Bcontroller%5D=News&tx_news_pi1%5Bnews%5D=4305&cHash=9fbac31598ab6ab106b39a45a4ffd8d4> under ongoing development, accessed 22 September 2022.
[63] Michael Cepic, Mariana Risetto, Mapping the European Legal Framework on Security Requirements for Cloud Computing Infrastructures in the Healthcare Sector, EDPL vol. 4, 2020

Furthermore, this section also showcases how the TVB-Cloud addresses the requirements imposed by these frameworks and draws conclusions what a party offering cloud services shall take into consideration.

2.2.1. Data Protection Legal Framework for Cloud Computing

In the field of data protection, the most important aspect which had to be considered was whether the project will process personal data in the cloud and what legal acts apply in that regard.

In the cross-border, European projects, which employ personal data in their cloud activities, one of the main positions of the frameworks, which have to be respected, takes the GDPR, aiming for the protection of natural persons. The GDPR and similar national laws are legal frameworks that protect personal information by imposing restrictions to storing, sharing, and processing of personal data. The cloud services such as those provided by the VBC project, are based on processing and storing different types of personal data, primarily the data such as neuroimaging data, derivatives thereof, neurosimulation results and associated metadata. The purpose of processing of those data is to enable personalized brain simulation.

First of all, in order to evaluate the GDPR's applicability to the project/cloud, it has to be assessed what types of data are foreseen to be stored and processed in the cloud. The definition of personal data and the definitions of categories of data which are considered personal as per Art.s 13, 14 and 15 GDPRhave to be taken into consideration. In the case of the TVB-Cloud, relevant definitions are "data concerning health"[64], which represents special categories of data, requiring higher level of protection defined in the GDPR.[65] To enable processing such data in the cloud, an appropriate legal basis has to be identified among the legal bases from Art. 6 and Art. 9 GDPR by the data controller as well as appropriate security measures have to be applied. Nevertheless, apart from the GDPR itself, it must be taken into account that GDPR leaves to the Member States the freedom to 'maintain or introduce further conditions, including limitations, with regard to the processing of genetic data, biometric data or data concerning health'[66], hence additionally requirements defined in national law have to be observed.

Another aspect which has to be determined in the context of processing personal data in the cloud are the roles assumed in the processing of personal data, which is related to the responsibility to implement and maintain the security standards imposed by the GDPR and determines how is responsible for the

[64] GDPR, Art. 4 (15), 'data concerning health' means personal data related to the physical or mental health of a natural person, including the provision of health care services, which reveal information about his or her health status.
[65] GDPR, Recital 53.
[66] GDPR, Art. 9(4).

compliance with the GDPR principles set out in Art.. 5 GDPR. The GDPR indicates two main roles which can be assigned to natural or legal entities participating – controller[67] and processor[68].

In a cloud, the role of a controller or processor is determined based on the type of interaction of involved players, such as Cloud Service Providers or other involved actors, such as the cloud storage provider and cloud service customer. The assumption of the mentioned roles in the cloud environment depends on the cloud structure and requires case-by-case assessment.

> 'If a provider is a data processor, a customer of cloud services is generally considered a data controller of the data stored in the provider's servers. This is typically the case for IaaS and PaaS services in which, in principle, the customer determines how the data is processed and the purpose for which it is processed. The customer is a processor if she or he is merely processing the personal data according to the wishes of a third party. This is typically the case of SaaS. GDPR assigns the responsibility for violations in the processing of personal data mainly to the cloud service customer, as a data controller, but adds a shared responsibility with the processor as joint controller when the customer does not have direct control of the data and its process'[69].

In the TVB-Cloud, the roles have been assessed and appropriately assigned. It has been established that when a user utilises TVB to process personal data, the user will always be a data controller, while the TVB-Cloud provider, as a service provider, will always be a data processor. It is because the user is conducting and steering the processing the processing through its interaction with the offered service, whereas the TVB is simply executing the instructions provided by the user. A user is in charge of control over the data hosted and processed with the TVB-Cloud and can independently or jointly determine the means of the data processing. What is also of importance is that the user can decide to stop processing and remove the data from the cloud at any time. The TVB provider is responsible to ascertain appropriate technical and organizational measures to protect the data provided by the data controller, i.e. for protecting the infrastructure with appropriately documented procedures and services employed by the users and on behalf of them. The user is required to accept terms of the TVB-Cloud use, that indicate the user's personal responsibility with regards to the GDPR compliance, including, inter alia,

[67] GDPR, Art. 4 (7) defines 'controller' as 'the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data'
[68] GDPR, Art. 4 (8), a processor is 'a natural or legal person, public authority, agency or any other body which processes personal data on behalf of the controller'. The WP 29 has set out two basic conditions for an entity to qualify as a data processor, these being a 'separate legal entity with respect to the controller' and 'processing personal data on his (the controller's) behalf'. WP29, Opinion 1/2010 (n 11) 1.
[69] Russo et al., 'Cloud Computing and the New EU General Data Protection Regulation' (2018) 5 IEEE Cloud Computing 1, 58-68.

security precautions, access permissions, personal responsibilities, monitoring, logging[70]. Data controllers must confirm they understand that active measures have been provided to protect sensitive data, but it is not excluded that certain vulnerabilities may still exist and a remaining risk for data protection incidents cannot be eliminated as it is an inherent risk of shared networks and computing systems. Additionally, as stated above, the GDPR requires the data controller to indicate a legal basis for processing as defined in Art. 6(1)(a) GDPR for processing to be considered lawful. In the case of the processing of 'special categories' of personal data, such as data concerning health, there must also be an exception applicable, as set out in Art. 9(2) GDPR, lifting the prohibition for processing being a general rule defined in Art. 9(1) GDPR. It is often the case in scientific research that in accordance with Art. 9(2)(a) GDPR, data subjects give their explicit consent to the processing of their personal data for one or more specific purposes. However, another exception may also be appropriate such as Art. 9(2)(j) GDPR on processing for the purposes of scientific research. In that case the national law must be considered additionally and the safeguards set out in Art. 89(1) GDPR have to be applied. The data controller has to ensure compliance with the data minimization principle, i.e. that the data uploaded to the TVB-Cloud is limited to only those needed for the purpose of the specific processing operation (GDPR Art. 5(1)(c)). Moreover, data controllers have to pseudonymise data as far as it does not compromise the research objectives and delete metadata that could lead to (re-)identification such as names, birth dates, behavioral scores, etc. The Usage Agreement provided to the data controllers on the TVB-Cloud defines that the data is temporarily stored and processed on servers to which different or additional areas of jurisdiction may apply. If the storage locations (countries and regions) are made clear, and the applicability of national and international laws has to be considered. Additionally, users acknowledge that the TVB-Cloud provides the services "as is" without claiming or guaranteeing correctness, accuracy, reliability, completeness, fitness, or usefulness for any purpose, reason, under any circumstance. Moreover, the user agrees that TVB on EBRAINS cloud providers, administrators or other users have no liability to any person or entity with respect to loss or damage caused directly or indirectly by the services provided by TVB on EBRAINS. Finally, Usage Agreement guarantees no involvement of subcontractors that are unable to guarantee that personal data will be processed under the conditions indicated in the Usage Agreement.

As the cloud-computing providers may usually be considered processors[71], they have to comply with the GDPR principles, defined in Art. 5 GDPR, whenever personal data comes into play. The security of

[70] Further information on terms of the TVB-Cloud are accessible at <http://www.ebrains.eu/terms> accessed 22 September 2022.

[71] According to Art. 32(1) GDPR, the controller and the processor are required to ensure the security of processing in an appropriate manner.

personal data must be ensured by application of appropriate technical and organisational measures (TOM's)[72], as required by Art. 5(1)(f) and Art. 32 GDPR.[73]

Art. 32(1) GDPR has to be considered already at the stage of selecting and setting up facilities and services, as corresponding to the principle of "data protection by design", defined in Art. 25 GDPR.

The first part of Art. 32(1) defines that the security measures applied should be appropriate to the risk, taking into account the „state of art". The state of art shall be understood in a dynamic and progressive manner – the controller (the user) and the processor (the cloud service provider) should adapt TOMs according to the technical developments.

Furthermore, the GDPR indicates which security measures are considered as appropriate:

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

First of all, in the process of assessing the appropriate level of security, the risks posed by processing of personal data have to be considered (Art. 32(2) GDPR). The risk component for the implementation of risk-based security measures needs to be evaluated through an objective assessment, to allow for determining whether data processing involves a risk or a high risk, taking into consideration the varying likelihood and severity for data subjects' rights and freedoms.[74] The type, scope, purposes and circumstances, i.e. context and characteristics of the processing, determine the relevance of the risk. Additionally, the risks indicated in Art. 32(2) GDPR, such as "accidental or unlawful destruction, loss, alteration, unauthorized disclosure of, or access to personal data transmitted, stored or otherwise processed" have to be considered. Further risks, which may show up in the course of conducting the summary of technical details, should be also taken as a basis for the risk assessment.

[72] Art. 5(1)f GDPR, 'Personal data shall be processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures ('integrity and confidentiality')'
[73] Johanna M. Hofmann and Alexander Roßnagel, ‚Rechtsverträgliche Gestaltung von Cloud Services' in Helmut Krcmar (ed.), Managmenet sicherer Cloud-Services (Springer Gabler 2018) 38.
[74] Michael Cepic, Mariana Risetto, 'Mapping the European Legal Framework on Security Requirements for Cloud Computing Infrastructures in the Healthcare Sector', EDPL vol. 4, 2020, p. 541; Hans-Jürgen Pollirer, 'Art 32 Sicherheit der Verarbeitung.' in Rainer Kyrim (edt.), *DatKomm* (2018) margin number 20 and 26.

Secondly, based on the risk assessment, appropriate TOMs have to be selected. Art. 32 GDPR provides an abstract list of the measures which have to be further defined to achieve compliance with privacy and data protection requirements. It should be emphasized that the GDPR does not lay down an absolute, static level of protection, but measures it against the individual risk. The wording of the GDPR, among others Art. 32 GDPR, is the reflection of the risk-based approach, and following that, the selection of TOMs must provide balance between the level of protection corresponding with the state of the art and risk evaluated. In selecting the measures, also implementation costs shall be taken into account.

The wording of Art. 32 GDPR does not include an enumerative catalogue of measures – quite the opposite; it allows for choosing other and/or additional measures that would enable demonstration of compliance with the legal obligations set out in Art. 32 GDPR.

In the context of encryption, sensitive data is encrypted before upload to the TVB-Cloud and remains encrypted unless a processing job is actively executed. It is a fundamental tool for data privacy which ensures that the data is unintelligible without decryption key.

Moreover, the TVB-Cloud encompasses the usage of public-key cryptography – the keys are created ad-hoc and independently in the case of each processing job and the design of the system does not allow any human to get into possession of the decryption key while the data is in the cloud. The only possibility to decrypt the data is at the site of their final processing, according to an automatic procedure.[75] The unencrypted data may exist only in envisaged sandboxes – isolated temporary memory locations with strict access rules.

Furthermore, since the TVB-Cloud architecture established three security zones based on the data content (Zone A, Zone B and Zone C), specific requirements have been established concerning which users will be allowed to access data stored in the different security zones. The administrator/s of the cloud platform have been identified and the specific access rights of such administrator/s determined.

By implementing all those measures, the TVB-Cloud addresses the requirements set out in the GDPR and applies data protection by design and by default approach (Art. 25 GDPR).

Another measure applied on the TVB-Cloud is apropriate definition and grading of the access rights, depending on the type of users utilising the platform (e.g. researchers, students, clinicians, patients). The specific access control mechanisms have been implemented, according to the legal advice indicated

[75] Michael Schirner et al., 'Brain simulation as a cloud service: The Virtual Brain on EBRAINS' (2022) NeuroImage, Volume 251, p. 8 <https://doi.org/10.1016/j.neuroimage.2022.118973>.

in D 2.1. The TVB web GUI direct access to systems (thevirtualbrain.apps.hbp.eu) where sensitive data are actively processed has been established. According to the rules of functioning of this system, users are required to log into the GUI with the credentials such as a secure password and cryptographic keys and further are able to access the data they uploaded or created, or the data that was made available to them through the role-based access control and permission management functionalities. The use of password and cryptographic key enables also secure delegated access for connecting different cloud services.

The conditions on the basis of which different users are allowed to gain access to the TVB-Cloud have been clearly identified. For example, since students will have access only to explorative functionalities of the platform, the information relating to what categories of data will be accessible through these limited functionalities has been provided.

According to Art.32(3) GDPR, an additional measure, which (in conjunction with the measures taken according to Arti. 32(1) GDPR), allows for achieving an appropriate level of protection, is compliance with approved codes of conduct (defined in Art. 40 GDPR) or approved certification procedure (defined in Art. 42 GDPR).

As of now, there are two codes of conduct applicable to cloud service providers adopted at the European level: “EU Data Protection Code of Conduct for Cloud Service Providers”, created by EU Cloud COC - Scope Europe and adopted by Belgian Data Protection Authority, and “Data Protection Code of Conduct for Cloud Infrastructure Services Providers (IaaS)”, developed by CISPE and adopted by French DPA.

With regard to the EU Cloud COC, given the broad definition of research infrastructure and the broad personal scope of this particular CoC (all cloud service providers are welcome to adopt the CoC), there is a window for research infrastructure owners built as cloud services to adhere to this CoC.

Regarding the second one, for scientific research infrastructures, the code could be of relevance if the research infrastructure providers are thought to develop and provide a cloud infrastructure service; however, the material scope of such code would not apply as the service could still be in development. This leads to the question whether ongoing research infrastructures, not fully developed yet, could be subject to such a self-assessment procedure as defined in this code.

Another measure which could be considered by cloud service providers is an application for a data protetction certification scheme. The data controller may choose whether they wish to be certified, and

whether such certification should be conducted in the framework of the competent supervisory authority or an accredited certification body. The application for such a certification is voluntary.

The project has been exploring the Codes of Conducts from a theoretical perspective, considering which one of them could be potentially relevant for the scientific research infrastructure. It is not excluded that the TVB could take them into account and adhere to a selected one in the future.

2.2.2.Legal framework for the Healthcare Sector

The healthcare sector is regulated by several legal acts at the European level, such as the Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices[76], and the Directive 2011/24/EU of the European Parliament and of the Council of 9 March 2011 on the application of patients' rights in cross-border healthcare[77]..

Nevertheless, none of them addresses the use of cloud services in the field and requirements related to it. It is however not excluded that the national regulations are here more detailed and do set out security requirements which have to be considered and implemented by the cloud providers.

There are also some soft law solutions, such as ICT standards which could be hypothetically issued for the ICT security certification in the healthcare area.[78]

2.2.3.Cybersecurity Legal Framework

The second legal framework significant for development and application of cloud services in a safe and secure manner considers the information security and at the European level is developed mostly in two major legal acts: NIS Directive [79]and the European Cybersecurity Act[80].

[76] Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC (Text with EEA relevance).
[77] Directive 2011/24/EU of the European Parliament and of the Council of 9 March 2011 on the application of patients' rights in cross-border healthcare.
[78] ENISA, *ICT security certification opportunities in the healthcare sector* (2019). This document covers functional requirements for a potential ICT security certification scheme for a widely understood healthcare sector, and summarises a common high-level functional security requirements for healthcare sector; or Cyber Security Requirements for Network Connected Medical Devices: This document 'summarises best practices for manufacturers of network-connected medical devices. These recommendations accompany regulatory requirements and are intended to support implementation and maintenance at an appropriate level of cyber security according to the current state of the art'.
[79] Directive (EU) 2016/1148 of the European Parliament and of the Council of 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union.
[80] Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013 (Cybersecurity Act) (Text with EEA relevance).

The NIS Directive's main objective is to introduce measures with a view to achieving a high common level of security of network and information systems within the Union to improve the functioning of the internal market (Art. 1 NIS Directive). Being a directive, it had to be transposed by the Member States into their national legal systems. The Member States were requested to create a national strategy on the security of network and information system and provide obligations for operators of essential services and DSPs – digital service providers. The cloud computing services were recognized as one of the DSP types (Art. 4 (5) and Annex III NIS Directive).

The NIS Directive applies to the legal entities providing cloud services, defined as 'a digital service that enables access to a scalable and elastic pool of shareable computing resources' (Art.. 4(19) NIS Directive).

According to Art. 16 (1) NIS Directive, the cloud service providers are required to identify and take appropriate and proportionate technical and organizational measures to manage risks created by the security systems. In line with the mentioned Art., the level of security which should be ascertained has to be appropriate to the risks posed, and should take into account the following elements:

a. the security of systems and facilities;
b. incident handling;
c. business continuity management;
d. monitoring, auditing and testing; and
d. compliance with international standards.

These elements have been further interpreted in the European Commission Implementing Regulation.[81]

Another obligation imposed on the cloud services provider in that regard is to prevent and minimize the impact of incidents affecting their network. According to the implementing regulation, 'with regard to incident handling referred to in point (b) of Art. 16(1) of Directive (EU) 2016/1148, the measures taken by the digital service provider shall include: (a) detection processes and procedures maintained and tested to ensure timely and adequate awareness of anomalous events; (b) processes and policies on reporting incidents and identifying weaknesses and vulnerabilities in their information systems; (c) a response in accordance with established procedures and reporting the results of the measure taken; (d) an assessment of the incident's severity, documenting knowledge from incident analysis and collection

[81] Commission Implementing Regulation (EU) 2018/151 of 30 January 2018 laying down rules for application of Directive (EU) 2016/1148 of the European Parliament and of the Council as regards further specification of the elements to be taken into account by digital service providers for managing the risks posed to the security of network and information systems and of the parameters for determining whether an incident has a substantial impact. L 26/48.

of relevant information which may serve as evidence and support a continuous improvement process.' (Art. 2(1) Implementing Regulation).

Moreover, it is required that the cloud service provider ensures appropriate business continuity management, i.e.

> "the capability of an organisation to maintain or as appropriate restore the delivery of services at acceptable predefined levels following a disruptive incident and shall include: (a) the establishment and the use of contingency plans based on a business impact analysis for ensuring the continuity of the services provided by digital service providers which shall be assessed and tested on a regular basis for example, through exercises; (b) disaster recovery capabilities which shall be assessed and tested on a regular basis for example, through exercises." (Art. 2(3) Implementing Regulation).

To fulfil the obligation of monitoring, auditing and testing, the Implementing Regulation indicates that there should be policies implemented which will define supervision, inspection and verification procedures. (Art. 2(4) Implementing Regulation).

The compliance with all above described points should be appropriately documented.

The NIS Directive establishes a framework according to which the Member States are required to:

- ensure that appropriate and proportionate Technical and organizational measures are defined and adopted by the DSPs,
- ensure that the DSPs act in a manner that prevents and minimizes the impact of incidents in case they occur,
- notify a competent authority or incident response team of any incident that has a substantial impact on the provision of cloud computing services without undue delay.

Additional, more relevant guidelines which allow for better understanding of the NIS Directive and provide recommendations and guidelines to the DSPs (so also to the cloud services providers) are introduced by ENISA.[82]

As stated in the section on data protection legal framework, the TVB-Cloud adheres to the NIS Directive requirements by providing the security solutions. They are the ones discussed in the description of Art.

[82] ENISA, *Technical Guidelines for the implementation of minimum security measures for Digital Service Providers* (2016).

32 GDPR requirements, to which a relevant technical documentation has been developed and maintained in the project, and the security measures will be continuously monitored and audited.

Another Europeanlegal instrument having impact on the cloud computing services security is the *Regulation (EU) 2019/881 of the European Parliament and of the Council of17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification, which entered into force in 27 June 2019* (EU Cybersecurity Act). This act, apart from defining the role of ENISA, lays down the EU cybersecurity certification framework and the procedure for the creation of EU cybersecurity certification schemes, including the ICT products, services and processes.[83] The scheme is currently under development under the leadership of ENISA.

Since the schemes are not ready yet, the TVB-Cloud could not adhere to them. This is left up tot he future decisions whether to adopt any kind of certification of this kind when it is available.

2.2.4.Cloud Computing Legal Framework

Based on the above analysis we draw a conclusion that whereas there is no unified and harmonized EU legal framework that applies to cloud computing as per the time of writing this deliverbale, the applicable framework is composed of various areas of law, such as data protection, consumer protection, cybersecurity environmental protection or intellectual property.

It should be, however, stressed that there are several legal instruments which, additionally to the two above mentioned, aim to, according to the European Cloud Strategy[84], regulate this field from different angles:

a. the Regulation (EU) 2018/1807 on the free flow of non-personal data – in parallel to the GDPR it increases legal certainty for the cloud service providers as it allows for free movement of the data in the European Union and raises trust by through the ongoing self-regulatory work on cloud switching and cloud security, conducted by the Digital Single Market Cloud Stakeholders Groups.[85]

[83]ENISA, EU Cybersecurity Certification Framework <https://www.enisa.europa.eu/topics/standards/certification> accessed 22 September 2022.
[84] European Commission, European Commission Cloud Strategy <https://ec.europa.eu/info/publications/european-commission-cloud-strategy_en>, accessed 22 September 2022.
[85] Michael Cepic, Mariana Risetto, 'Mapping the European Legal Framework on Security Requirements for Cloud Computing Infrastructures in the Healthcare Sector' (2020) EDPL vol. 4, 546.

b. ENISA is currently developing a single European cybersecurity certification scheme for cloud services[86].
c. The development of Codes of Conduct for cloud service providers, defining the data protection in the cloud, is strongly encouraged by the European Commission. There are already two Codes of Conduct adopted at the European level.

There are several soft law instruments and recommendations applicable to the security which may be considered by cloud service providers:

a. ISO standards, such as ISO 27001, ISO/IEC 27017:2015 [ISO/IEC 27017:2015] Information technology — Security techniques — Code of practice for information security controls based on ISO/IEC 27002 for cloud services; or BSI. As indicated in the replies to the survey included in Annex I, both standards have been adopted in the context of the TVB-Cloud.
b. European Strategies on cloud computing:
    a. European Cloud Initiative – Building a competitive data and knowledge economy in Europe. This initiative focuses on the scientific community and computing capacity and on extending the capabilities to other areas of life. It entails:
        i. the development of trusted, open environment for the use by the scientific community to safely and securely store, share and re-use scientific data and results: the European Open Science Cloud (EOSC)
        ii. the deployment of the super-computing capacity, fast connectivity and high-capacity cloud solutions in the context of the new European Data Infrastructure (EDI), supporting EOSC.

The VBC adheres to the requirements defined by the NIS Directive as well as adheres to the ISO Standard concerning information security, as it has been indicated in the answers to the questionnaires included in Annex I to this Deliverable.

2.2.5. Conclusion

Considering the the analysis provided aboveon the legal framework that governs the privacy and security aspects to be considered by the cloud providers, there are several lessons learned and recommendations which the TVB-Cloud project provides.

[86] ENISA, Cloud Certification Scheme: Building Trusted Cloud Services Across Europe <https://www.enisa.europa.eu/news/enisa-news/cloud-certification-scheme>, accessed 22 September 2022.

1. It is of great importance to understand legal and regulatory frameworks as well as restrictions that apply to the information security and handling of the data (especially personal data) in the custody of the cloud service providers and the users, even before the design phase.
2. Cloud providers shall disclose the terms and conditions that apply to the the services provided by them, as well as shall clearly describe the policies governing the cloud provision.
3. Cloud providers shall thoroughly plan and implement appropriate security measures and constantly monitor, test and audit them, as well as keep the documentation describing those measures in a transparent and clear manner.

The TVB-Cloud took into consideration the described frameworks and applied solutions that comply with the requirements established therein, as it has been presented above, and can be reflected in D7. 2 (ANNEX TO D7.2 ON THE LEGAL FRAMEWORK APPLICABLE TO CLOUD COMPUTING – FOCUS ON SECURITY REQUIREMENTS ('TECHNICAL SECURITY MEASURES').

2.3. Prospective legal framework

This sub-chapter outlines selected legal aspects that might affect the TVB-C in the future.

2.3.1. EU Proposal Regulation on Artificial Intelligence

Here we provide an analysis of the applicability of the Proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on Artificial Intelligence (Artificial Intelligence Act) (hereinafter, "AIA")[87] to the **VirtualBrainCloud software**[88]. With this aim, such a proposal is contextualized, its material scope analyzed, and the obligations arising out of such legal instruments addressed.

The Proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on Artificial Intelligence (Artificial Intelligence Act) (hereinafter, "AIA") and amending certain union legislative acts, was published by the Commission on 21 April 2021[89]. It is the materialization of one of the components (Artificial intelligence) identified in the **Digital Compass**, which sets out the European Union's concrete **digital ambitions for 2030**[90].

It is important to note that at the time of writing this deliverable (D2.5), this legal instrument has a proposal nature; therefore, this section addresses a preliminary analysis on a possible applicability of

[87] Available at <>, accessed 22 September 2022.
[88] See Task 6.4 Annex I Part B GA.
[89] Available at <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206>, accessed 22 September 2022.
[90] European Commission, 'Shaping Europe's digital future. Europe's Digital Decade' <https://digital-strategy.ec.europa.eu/en/policies/europes-digital-decade>, accessed 22 September 2022.

the AIA to the TVB-Cloud AI software as of the status quo of the Proposal. The analysis performed below may change if the proposal is modified thereafter. As reported in February 2022, the EU legislative train shows that the instrument is now (September 2022) being discussed by the co-legislators, the European Parliament and the Council (EU Member states). In Council, negotiations to find a common position between Member states have started. In November 2021, the Slovenian presidency presented a progress report (draft compromise) on discussions held so far within the Council on the AI draft proposal[9192].

Analysis AIA Applicability

The AIA pursues a number of reasons of public interest, such as guarantee a high level of protection of health, safety and fundamental rights, and ensure the free movement of AI-based goods and services cross-border, thus preventing Member States from imposing restrictions on the development, marketing and use of AI systems, unless explicitly authorised by this Regulation.

The proposed Regulation takes a risk-based approach; classifying different AI software based on the likelihood they may pose a significant risk to the health and safety of natural persons.

It must be reminded that the AIA is designed to 'fit into a rather sophisticated system of existing laws many of which will not explicitly address AI systems, but will, without any doubt, capture a wide range of activities that make use of AI systems'[93]; in particular the relation with the Medical Device Regulation (see 'Categorization' below).

Personal scope

As per the personal scope of the AIA, Art. 2(1) sets forth that the regulation applies to:

(a) providers placing on the market or putting into service AI systems in the Union, irrespective of whether those providers are established within the Union or in a third country;

(b) users of AI systems located within the Union;

[91] European Parliament, 'Proposal for a Regulation on a European approach for Artificial Intelligence In "A Europe Fit for the Digital Age' <https://www.europarl.europa.eu/legislative-train/theme-a-europe-fit-for-the-digital-age/file-regulation-on-artificial-intelligence> accessed 22 September 2022.

[92] Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - Progress report - Interinstitutional File: 2021/0106(COD), 22 November 2021, <https://data.consilium.europa.eu/doc/document/ST-13802-2021-REV-1/en/pdf>, accessed 29 November 2022.

[93] Christiane Wendehorst, 'The Proposal for an Artificial Intelligence Act COM(2021) 206 from a Consumer Policy Perspective', Austrian Federal Ministry of Social Affairs, Health, Care and Consumer Protection, 14 December 2021.

(c) providers and users of AI systems that are located in a third country, where the output produced by the system is used in the Union.

Taking into consideration the TVB-Cloud constellation, the consortium may fall within the first group of entities.

Material Scope

As per the material scope, Art. 3(1) definition of AI system entails two components:

1. a 'software developed with one or more of the following techniques and approaches listed in Annex I':

- Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning;
- Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems;
- Statistical approaches, Bayesian estimation, search and optimization methods;

This catalogue is not exhaustive, as the Commission may amend this list in order to update it to market and technological developments, and therefore, partners should follow the legisöative developments closely.

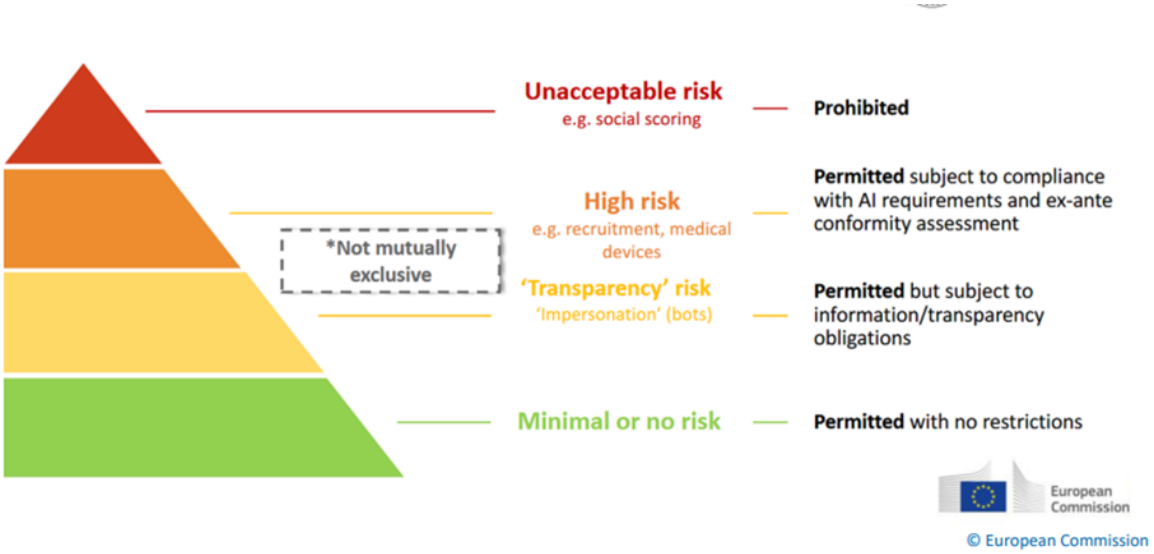
2. This software can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with.

Both components have to be fulfilled. In the case of TVB-Cloud, such machine learning techniques are/were developed and are aimed to be incorporated on the TVB-Cloud solution as part of Task 6.4 (Interactive interfaces for interventions, diagnostics and prognostics (M01-M46)) and tasks in WP8 (Personalised Simulation). Additionally, in order to determine the scope of applicability, it does not **appear that AI systems produced by the TVB-Cloud would fall within the prohibited AI systems as described in Art. 4 of the AIA**[94].

[94] See Art. 5 AIA

2.3.2.Categorization

As stated above, the next step is to evaluate if the TVB-Cloud system are categorized under a low, medium or high-risk AI system in order to determine the obligations attached to the system providers.



For this reason, High-risk AI systems are defined in Art. 6.1 as those systems which are intended to be used as a **safety component of a product** or **is itself a product covered by the Union legislation**. Art. 6(2) AI systems under such category are those included in the following areas, e.g. **Biometric identification**, or **Management and operation of critical infrastructure** (road traffic and the supply of water, gas, heating and electricity), Education and vocational training, among others. Additionally, future extensions of Annex III according to criteria listed in Art. 7.

- AI system deploying subliminal techniques beyond a person's consciousness in order to materially distort a person's behaviour in a manner that causes or is likely to cause that person or another person physical or psychological harm;
- AI system exploiting any of the vulnerabilities of a specific group of persons due to their age, physical or mental disability, in order to materially distort the behaviour of a person pertaining to that group in a manner that causes or is likely to cause that person or another person physical or psychological harm
- AI systems by public authorities for the evaluation or classification of the trustworthiness of natural persons over a certain period of time based on their social behaviour or known or predicted personal or personality characteristics, with the social score
- Use of "real-time" remote biometric identification systems in publicly accessible spaces for the purpose of law enforcement unless it is used for selected reasons (further defined in the AIA).

Derived from this Art., it must be evaluated whether the TVB-Cloud AI software **is itself a product** covered by the Union legislation, in the meaning of Art. 6(2). Ruling out the requirement of 'as a **safety component of a product**', it needs to be evaluated whether the TVB-Cloud AI software **is itself a product covered by the Union legislation.** In this regard, a relevant legal instrument of possible applicability is the **Medical Device Regulation**[95] (hereinafter, 'MDR').

This Regulation applies to a "medical device" defined as 'any instrument, apparatus, appliance, **software**, implant, reagent, material or other Art. intended by the manufacturer to be used, alone or in combination, for human beings for one or more of the following specific medical purposes: - diagnosis, prevention, monitoring, prediction, prognosis, treatment or alleviation of disease'[96]. Given the fact that the TVB-develops and validates a decision support system that provides access to high quality multi-disciplinary data for clinical practice, materialized in a cloud-based brain simulation platform to support personalized diagnostics and treatments in NDD, it seems that such regulation might apply to the TVB-Cloud end-product.

Therefore, many medical devices will qualify as "high-risk" under the proposed regulation – assuming they must undergo a third-party conformity assessment[97]. That is, TVB-Cloud anticipated device is qualified herein as a high-risk AI device under the proposed AIA.

Last but not least, the consortium should keep of monitoring if Annex III may include future extensions according to criteria listed in Art. 7 is extended in the future.

High-risk System Obligations

Under the proposed AIA, high risk AI is subject to a number of obligations. These obligations are addressed to the provider of a High-risk AI system, which include the obligations delineated below. The AIA also delineates obligations for the user and importer of such systems:

- A risk management system (AIA, Art. 9)

Requires AI providers to establish and maintain a continuous, iterative risk management system through the AI system's lifetime. The risk management system must include an analysis of the foreseeable risks

[95] Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC.
[96] MDR, Art. 2(1).
[97] See AIA Art. 6 (defining a high risk system as one in which: "the AI system is intended to be used as a safety component of a product, or is itself a product, covered by the Union harmonisation legislation listed in Annex II [which includes the MDR and IVDR at Annex II, Section A, 11. & 12.]" and in which "the product whose safety component is the AI system, or the AI system itself as a product, is required to undergo a third-party conformity assessment with a view to the placing on the market…"

associated with the AI system; an estimation and evaluation of risks that may emerge; an evaluation of other risks based on post market analysis (Art. 61); and risk management measures. Risk mitigation measures should include elimination or reduction of risk through adequate design and development and a testing procedure to be performed during the development process in order to identify the most appropriate risk management measures. Residual risks must be communicated to users.

- Data governance and management requirements (Art. 10)

Requires AI systems to be trained and validated on data sets that meet specific quality criteria. For example, "[t]raining, validation and testing data sets shall be relevant, representative, free of errors and complete".[98] The training and validation and testing data sets needs to also follow an appropriate data and management practices. Relevant for the TVB-Cloud is paragraph 5 of Art. 10, which states that providers of high-risk AI systems may process special categories of personal data referred to in Art. 9(1) of Regulation (EU) 2016/679, Art. 10 of Directive (EU) 2016/680 and Art. 10(1) of Regulation (EU) 2018/1725 to the extent that is necessary for the purposes of ensuring bias monitories and correction in relation to the systems and are subject to appropriate safeguards for the fundamental rights and freedoms of natural persons, including technical limitations on the re-use and use of state-of-the-art security and privacy-preserving measures, such as pseudonymisation, or encryption where anonymisation may significantly affect the purpose pursued.

- Technical documentation (Art. 11)

High-risk AI system providers must provide technical documentation that demonstrates the AI system complies with the AIA and sets forth "all the necessary information to assess the compliance of the AI system with those requirements".[99] The Commission has provided a list, as part of Annex IV to the proposed legislation, which sets out specific details such a technical documentation must include. Where the MDR, as addressed above is applicable, the technical documentation/application under the MDR might suffice to meet this requirement.[100]

- Recording system/Automatic logging system (Art. 12 and 20)

The AI system must have the ability to automatically record events while the system is operating, enabling the "monitoring of the operation of the high-risk AI system with respect to the occurrence of situations that may result in the AI system presenting a risk" to the health and safety (or fundamental

[98] AIA, Art. 10(3).
[99] AIA, Art. 11.
[100] AIA, Art. 11(2)

rights) of persons "or lead to a substantial modification, and facilitate [] post-market monitoring."[101] For the logs to be automatically generated, such logs must be under the control of the providers via a contractual arrangement with the user or stipulated by law.

- Transparency requirements and information to users (Art. 13)

Requires AI systems to be developed so that their "operation is sufficiently transparent to enable users to interpret the system's output and use it appropriately."[102] Transparency measures should include a set of instructions for use (which include, among other things details regarding the system's performance, the level of accuracy, robustness and cybersecurity for which the device has been tested, warnings regarding risks of normal and foreseeable misuse, and specifications regarding input data); pre-determined possible changes to the device and its performance; human oversight measures (Art. 14); and expected lifetime of the device and necessary maintenance measures.

- Human oversight (Art. 14)

Requires AI systems to be designed and developed so that they can be effectively overseen by a natural person with the aim to "preventing or minimising the risks to health, safety or fundamental rights".[103] Specific capabilities of the system, as appropriate, should include enabling the users of the system to: (1) understand the capacities and limitations of the system and monitor its operation; (2) remain aware of the possible tendency to over-rely on output produced by an AI system particularly for recommendations for decisions to be taken by natural persons; (3) correctly interpret the output; (4) to be able to decide not to use the AI system or otherwise disregard, override or reverse the output of the system; (5) be able to intervene on the operation of the AI system or interrupt the system through a stop button or similar procedure.

- Accuracy, robustness, and cybersecurity (Art. 15)

Requires AI systems to exhibit an "appropriate level of accuracy, robustness and cybersecurity".[104] The levels of accuracy and relevant accuracy metrics must be provided in the accompanying instructions of use. Accuracy indicates the device's ability to be "resilient as regards errors, faults or inconsistencies that may occur within the system"; robustness may be achieved through technical redundancy solutions. The system should be able to withstand attempts by unauthorised third parties to alter the

[101] AIA, Art. 12
[102] AIA, Art. 13.
[103] AIA, Art. 14.
[104] AIA, Art. 15.

use or performance of the system by exploiting its vulnerabilities, therefore guaranteeing a level of cybersecurity.

- Quality management systems (Art.s 16 and 17)

A quality management system must be put in place, documented in the form of written policies, procedures and instructions, including a long list of items specified by the Art.[105] For example, "a strategy for…compliance with conformity assessment procedures and procedures for management of modifications to the high-risk AI system;"[106] "techniques, procedures and systematic actions to be used for the design, design control and design verification of the high-risk AI system;"[107] and "systems and procedures for data management, including data collection, data analysis, data labelling, data storage, data filtration, data mining, data aggregation, data retention and any other operation regarding the data that is performed before and for the purposes of the placing on the market or putting into service of the high-risk AI system."[108]

- Conformity assessment procedure and declaration of conformity/registration (Art.s 19, 40, 41, 43 & 44, 48 and 51)

A conformity assessment procedure is required under Art. 43 for AI devices. In particular, for medical devices are able to rely on the conformity assessment procedure under the MDR and therefore, satisfy this requirement as long as the requirements of Chapter 2 of the AIA and AIA's Annex VII, Points 4.3-4.6 are part of that conformity assessment procedure.

The provider must draw up a declaration of conformity, stating that system's compliance with Chapter 2 of the AIA for each system.[109] For medical devices this declaration will also include the declaration of conformity with the MDR. Also, the system must be registered in an EU public database.[110]

- Duty of information/ cooperation with competent authorities (Art.s 22 and 23)

Where a risk is posed on the product that might after the health and safety of a natural person in accordance with Regulation (EU) 2019/1020[111], this fact needs to be informed to the national

[105] AIA, Art. 17.
[106] Ibid.
[107] Ibid.
[108] Ibid.
[109] AIA, Art. 48.
[110] AIA, Art.s 51, 60.
[111] Regulation (EU) 2019/1020 of the European Parliament and of the Council of 20 June 2019 on market surveillance and compliance of products and amending Directive 2004/42/EC and Regulations (EC) No 765/2008 and (EU) No 305/2011 (Text with EEA relevance).

authorities. Upon request of national authorities, providers of such systems must provide necessary information on the obligations described above.

- Post marketing monitoring obligations (Art. 61)

A post-market monitoring system for the device must be established and documented by AI providers. In the case of medical devices, post-market monitoring system under the MDR qualify as a post-market monitoring system under the AIA, if the requirements in paragraphs 1-3 of Art. 61 are also integrated into the system. Therefore, the duplication of obligations under two different regimes are avoided.

The AIA suggests measures to support innovation by offering the building of "regulatory sandboxes" made to test "innovative AI systems".[112] One consequence of such structure is that it may allow for some flexibility with respect to the (non-)existence of an independent legal basis under the GDPR for personal data used to test AI devices making use of the sandbox. This might not apply to TVB-Cloud but should be considered if the system, as according to the TVB-Cloud business plan (D9.2 Finalised IPR and Exploitation plan (business plan version 1)) from where it can be stated that the system will not furthered on new AI systems.

2.3.3.Non-compliance

Non-compliance with the AIA regulation obligations leads to severe penalties. For example, the non-compliance with data and data governance requirements face fines of up to 30 million Euro or 6% of the total worldwide annual turnover for the preceding financial year, whichever is higher, could apply. In the event of non-compliance with other AIA requirements, fines of up to 20 million Euro, or 4% of the annual turnover could apply.[113]

2.3.4.Art.Relevant Opinions on the AIA proposal

Since the AIA proposal was realised by the Commission, several opinions/ progress reports of the EU assigned committees/organs/entities involved in the legislative procedure have been issued. Such opinions have an impact on the scope of the AIA, and therefore a having a brief overview of possible modifications to the current proposal which are relevant to TVB-Cloud might help understanding expectations and possible modifications. Suggestions/opinions relevant for the TVB-Cloud are highlighted.

[112] AIA, Art.s 53, 54.
[113] AIA, Art. 71.

- **Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - Progress report- President of the Council – 22 November 2022**

The main areas addressed by the Slovenian Presidency of the Council in the partial compromise are presented below:

- (Scope of the AIA) - The AIA should not apply to AI systems and their outputs used for the sole purpose of research and development.
- (Definitions) In particular, the definition of an AI system was stated that needs to be 'circumscribed to ensure more legal clarity and to better reflect what should be understood by an AI system for the purposes of the AIA, with an explicit reference indicating that any such system should be capable of determining how to achieve a given set of human defined objectives by learning, reasoning or modelling'. As the representative of Denmark stated in the public meeting[114], this change is intended to prevent the inclusion in the scope of the proposed regulation of more traditional software systems that are normally not considered as artificial intelligence like simple statistical systems. Level of autonomy should be consider as characteristics of an AI system. In relation to this modification, the list of techniques and approaches in Annex I has also been refined to ensure more clarity on which systems are covered by this definition (this not yet public).
- (classification of high risk AI systems) High-risk system should be the exception. The provisions defining the rules for classification of high-risk AI systems have been thoroughly revised to ensure better legal clarity and readability of the text. In relation to the conditions for the amendments to Annex III as in Art. 7, further updates have been made to the criteria for such amendments to ensure more legal clarity.
- Complex issues have been identified in particular Requirements for high-risk AI systems, responsibilities of various actors in the AI value chain which are now mostly carried by the provides should be distributed in the value chains; compliance and enforcement regarding penalties are excessive for SMEs and start-ups; and the relationships with existing legislation needs to be better connected as there are many overlapping obligations and it may be necessary to eliminate potential legal discerption.

- **Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)**

[114] Informal video conference of telecommunications ministers Public session Thursday, 14 October 2021 - 09:59. Available at <https://video.consilium.europa.eu/event/en/24933>, accessed 29 November 2022.

- (Risk-based approach) The EDPB and EDPS take note of the choice of providing an exhaustive list of high-risk AI systems, which in their opinion 'might create a black-and-white effect, with weak attraction capabilities of highly risky situations, undermining the overall risk-based approach underlying the Proposal'. In particular since 'the list in in annexes II and III of the Proposal lacks some types of use cases which involve significant risks, such as for assessing medical treatments or for health research purposes'. These two examples are relevant for the TVB-Cloud solution, and should be in any way considered as examples to be then included in Annexes II or II of the AIA.
- (High-risk AI systems) The recommendation aims to conduct ex ante a third-party conformity assessment.
- Parallelisms with GDPR:
    - Risk assessment by provider of the AI system (AIA), and user (AIA) as the controller of the AI system under GDPR.
    - Classification of 'high risk' under AIA does not trigger high-risk processing under GDPR. This entails that where the high-risk systems may be permissible but it does not imply that the associated processing of personal data is lawful.

Conclusion

The AIA in its current draft status is likely to apply to the TVB-Cloud end-product under the definition of Art. 3(1) of the AIA, and it is might be likely to be categorized under a high-risk AI system, which under the AIA poses strict obligations to the providers of such systems at the design and developing stage of the AI system.

2.3.5. Proposal for a European Health Data Space Regulation

When it comes to upcoming regulation in the context of data protection laws we need to furthermore highlight the proposal for a European Health Data Space Regulation[115] that is going to build a digital health ecosystem,[116] where inter alia

(a) patients and clinicians can access data throughout Europe in the form of electronic health records (primary use), and

[115] Proposal for a regulation of the European Parliament and of the Council on the European Health Data Space, COM/2022/197 final.
[116] European Commission, 'European Health Data Space' <https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en>, accessed 16 September 2022.

(b) researchers can access and use data for the benefit of society (secondary use).[117]

The proposed EDHS builds upon existing legislation, that is in particular the GDPR, the Medical Devices Regulation[118], the In Vitro Diagnostics Regulation[119], the proposed Artificial Intelligence Act[120], the proposed Data Governance Act[121], the proposed Data Act[122], the NIS Directive[123] and the CBHC Directive[124].[125] Due to its special nature with regard to the governance of the (re-)use of health data especially in the context of research, it is likely that the EHDS will be the most specific law governing such processing next to the GDPR; it can therefore reasonably be considered the lex specialis in the context of the aforementioned laws.

From a content perspective the proposed EHDS can basically be divided into two parts. Chapter 2 and 3 govern the use of health data for electronic health records. Chapter 3 regulates the secondary use of said data. As part of the proposed EHDS, Member States will have to set up so called (national) Health Data Access Bodies (HDAB)[126] through which health data will be shared under the EHDS. On the one hand research institutions, but also private companies such as pharmaceutical companies, that collect and process health data, will need to make these data available to the HDABs.[127] On the other hand researchers and others will be able to request such data from these institutions through the HDAB in an anonymised and, where applicable, pseudonymised format.[128]

Conclusion

[117] Proposal for a regulation of the European Parliament and of the Council on the European Health Data Space COM/2022/197, Recital 1 „*The aim of this Regulation is to establish the European Health Data Space ('EHDS') in order to improve access to and control by natural persons over their personal electronic health data in the context of healthcare (primary use of electronic health data), as well as for other purposes that would benefit the society such as research, innovation, policy-making, patient safety, personalised medicine, official statistics or regulatory activities (secondary use of electronic health data). In addition, the goal is to improve the functioning of the internal market by laying down a uniform legal framework in particular for the development, marketing and use of electronic health record systems ('EHR systems') in conformity with Union values.*".
[118] Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC (Text with EEA relevance.).
[119] Regulation (EU) 2017/746 of the European Parliament and of the Council of 5 April 2017 on in vitro diagnostic medical devices and repealing Directive 98/79/EC and Commission Decision 2010/227/EU (Text with EEA relevance.).
[120] Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) COM/2021/206 final.
[121] Proposal for a Regulation on European data governance (Data Governance Act) COM/2020/767 final.
[122] Proposal for a Regulation on harmonised rules on fair access and use of data (Data Act) COM/2022/068 final.
[123] Directive (EU) 2016/1148 concerning measures for a high common level of security of network and information systems across the Union.
[124] Directive 2011/24/EU on the application of patients' rights in cross-border healthcare.
[125] Proposal for a regulation of the European Parliament and of the Council on the European Health Data Space, COM/2022/197 final, 3.
[126] EHDS, Art. 36(1).
[127] EHDS, Art. 2(2)(y) in connection with Art. 44(1).
[128] EHDS, Art. 2(2)(z) in connection with Art. 45 ff.

The implications of this Proposal to the Project are two folded. In the case such proposal cames into force during the lifespan of the project, reseachers could request access to certain datasets needed for the project development. Otherwise, the TVB-Cloud would need to consider incuding a role in its data govervnace which would respond to the data –related requests arding out of the EDHS framework.[129]

3. Ethical framework and analysis

The Ethics Guidelines on Trustworthy AI, which were developed by the EU High-Level Expert Group (HLEG) on AI in 2019[130], has three central tenets:

1. AI should be lawful, complying with all applicable laws and regulations;

2. AI should be ethical, ensuring adherence to ethical principles and values; and

3. AI should be robust, both from a technical and social perspective, since, even with good intentions, AI systems can cause unintentional harm.

The Guidelines state that each of these three components is necessary but not sufficient in itself to achieve Trustworthy AI, highlighting the need for research initiatives such as TVB-Cloud to work towards embedding all three components in the development and deployment of AI systems.

The first section of this deliverable primarily focuses on the first tenet: compliance of TVB-Cloud with all applicable laws and regulations. In this section, we address the second tenet: ethical implications and adherence to ethical principles and values. We build on Deliverable 2.1, which identified a number of foundational ethical principles of health research that have particular relevance to the project. D2.1 also highlighted key issues relating to two distinct areas within the TVB-Cloud project: first, scientific research using human data (personal data); and second, the use of predictive computational software, also referred to in this report to as Artificial Intelligence (hereinafter, 'AI'), in health. Here, we revisit the key ethical issues for TVB-Cloud, highlighting applicable sections of formal ethical and legal frameworks that have been developed since the launch of TVB-Cloud in 2018, and discussing their relevance and applicability to the primary outputs, and future development, of the TVB-Cloud project.

3.1. Key ethical considerations for TVB-Cloud

[129] For the conditions on datasets, details on request see Chapter 3 Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the European Health Data Space.
[130] Ethics Guidelines for Trustworthy AI (2019), published by the independent High-Level Expert Group on AI.

Conducting ethical research in health projects such as TVB-Cloud requires that the research project in question meets certain fundamental ethical principles. The ethics of health research involving human participants requires a context-sensitive approach and a balanced risk assessment about likely harm to the individual (or his/her community) on the one hand, versus the right to perform research for societal benefit on the other hand.

In D2.1, four key areas of ethical considerations were identified and discussed: fairness, non-discrimination, informed consent, and confidentiality. In addition, D2.1 discussed the role of research ethics committees and the establishment of an external advisory board for TVB-Cloud on legal and ethical issues. In this section, we reflect on the ethical issues identified in the initial legal and ethical framework for TVB-Cloud, and where relevant, discuss how these issues were raised over the duration of the project, identifying ways to address them during future development of TVB-Cloud. Finally, we expand on this analysis and integrate the perspectives of a key stakeholder group through public involvement work with the European Dementia Carers Working Group.

3.1.1. Fairness and equity

In 'Principles of Biomedical Ethics' by Beauchamp and Childress, first published in 1979[131], four key ethical principles were identified: respect for autonomy, beneficence, non-maleficence and justice. Fairness and equity are closely linked to the principle of justice described in this publication. Equity is synonymous with fairness, and is a structural and systemic concept that is defined by the World Health Organisation as "*the absence of avoidable or remediable differences among groups of people, whether those groups are defined socially, economically, demographically or geographically*." In the setting of health research projects such as TVB-Cloud, equity and fairness extends to participation in research. Specifically, individuals or groups should not be excluded from participating in research studies based on gender, ethnicity or other factors unrelated to the scientific goal of the clinical research study – just as these groups should not be marginalised from health benefits due to these factors. To ensure equitable access to research participation, researchers are encouraged to identify and address barriers to participation, for example by embedding measures to enhance inclusion (e.g allowing people with cognitive impairments to be accompanied by supporters or caregivers), and ensuring that tools employed in studies are suitable for use by a broad range of groups and not culturally biased, and have been validated on the groups on which they are being used (e.g using neuropsychological tests that are less affected by cultural or socioeconomic factors, and verifying the lack of bias by validation studies with a diverse range of participants; Gove et al, 2021[132]). Researchers should also avoid dismissing

[131] Beauchamp and Childress (1979) Principles of Biomedical Ethics, published by Oxford University Press

[132] Gove D, Nielsen TR, Smits C, Plejert C, Akhlak Rauf M, Parveen S, Jaakson S, Golan-Semesh D, Lahav D, Kaur R, Herz MK, Monsees J, Thyrian JR, Georges J (2021). The challenges of achieving timely diagnosis and culturally

groups as "hard to reach", which risks placing the blame on the potential participants rather than structural factors, attitudes and assumptions that can lead to discrimination[133].

Equity, and fairness in particular, are also a feature of many ethical frameworks for AI. The Ethics Guidelines for Trustworthy AI specifies fairness as one of the four principles for lawful, ethical and robust AI (along with respect for human autonomy; prevention of harm; and explicability). Fairness is closely linked to the rights of EU citizens as enshrined in the foundational treaties, in particular the right to non-discrimination, solidarity, and justice. The Guidelines define how AI should respect the principle of fairness, stating that:

The development, deployment and use of AI systems must be fair (…) we believe that fairness has both a substantive and a procedural dimension. The substantive dimension implies a commitment to: ensuring equal and just distribution of both benefits and costs, and ensuring that individuals and groups are free from unfair bias, discrimination and stigmatisation.(….) Equal opportunity in terms of access to education, goods, services and technology should also be fostered. Moreover, the use of AI systems should never lead to people being deceived or unjustifiably impaired in their freedom of choice. Additionally, fairness implies that AI practitioners should respect the principle of proportionality between means and ends, and consider carefully how to balance competing interests and objectives. The procedural dimension of fairness entails the ability to contest and seek effective redress against decisions made by AI systems and by the humans operating them. In order to do so, the entity accountable for the decision must be identifiable, and the decision-making processes should be explicable." (Ethics Guidelines for Trustworthy AI, EU HLEG, April 2019)

The OECD AI principles[134], which were adopted shortly after the EU HLEG Guidelines were published, similarly references fairness as a foundational, values-based principle for ethical AI, stating that AI should be designed in a way that respects the rule of law, human rights, democratic values and diversity, including appropriate safeguards to ensure a fair and just society. It should be noted, however, that there is an ongoing debate in bioethics on the value of principlist approaches (the belief that a healthcare decision will be morally justified if it is consistent with relevant ethical principles and judgements) compared to more narrative approaches (where every moral situation is unique and

appropriate care of people with dementia from minority ethnic groups in Europe. Int. J. Geriatr.Psychiatry https://doi.org/10.1002/gps.5614

[133] Alzheimer Europe (2018) The development of intercultural care and support for people with dementia from minority ethnic groups (lead author: Dianne Gove).

[134] Recommendation of the Council on Artificial Intelligence (2019) and AI Principles of the Organisation for Economic Cooperation and Development/OECD: https://oecd.ai/en/ai-principles

healthcare decisions are justified if they 'fit' with the narrative of that individual's life)[135]. While principlist approaches provide a structured method of objectively supporting ethical decisions, it is sometimes argued that they risk overlooking issues that are of importance to individual patients; have a tendency to generalise; and are challenged in real-life situations where it is important to consider competing interests. On the other hand, while narrative approaches respect the unique and personal stories of individuals ("personalised ethics", akin to personalised medicine), there may be disparities in "whose stories are told, whose stories are heard, and whose stories are believed", based on personal biases or internalised stigma[136]. While the two approaches share few similarities in theory, there may, however, be more complementarity in practice; for example, principlist approaches can be enhanced using narrative skills such as empathetic listening and support to consider the uniqueness of individual situations.

As illustrated above, the principles of fairness and equity as framed in the EU and OECD frameworks have a degree of overlap with the principle of non-discrimination (addressed in greater detail in section 3.3 below). These frameworks identify non-discrimination as the primary component of fairness, through avoidance of biases (i.e. algorithmic biases that arise due to cohort bias, minority bias in training data, biases linked to interactions with clinicians or patients). However, ethicists have pointed out that technical methods to address algorithmic fairness do not always account for causal interactions between biological, environmental and social factors, stating that "*framing fairness as a purely technical problem solvable by the inclusion of more data or accurate computations is ethically problematic…*" and that "*relying on the so-called veneer of technical neutrality could exacerbate harms to vulnerable groups*." (McCradden et al, 2020[137]). To address these issues, Giovanola & Tiribelli (2022[138]) call for a more complex concept of fairness, argue that fairness requires more than non-discrimination and avoidance of bias, as it is also a socio-relational construct, requiring "*a commitment to ensure equal respect for persons as individuals*". In technical terms, this could involve the use of "*compensatory tools that mitigate social disparities…[machine learning algorithms] could be informed with sensitive traits, allowing them to evaluate who in a certain health domain may require compensatory tools if* ***explicitly*** *requested by the subject.*" In this concept of fairness, understanding and respect of individual needs and perspectives is viewed as key, requiring the considered implementation of AI innovations in healthcare

[135] McCarthy J (2003) Principlism or narrative ethics: must we choose between them? BMJ Medical Humanities http://dx.doi.org/10.1136/mh.29.2.65

[136] Saulnier KM (2020) Telling, hearing and believing: a critical analysis of narrative bioethics. J. Bioeth. Inq. https://doi.org/10.1007/s11673-020-09973-y

[137] McCradden M, Joshi S, Mawzi M and Anderson JA (2020) Ethical limitations of algorithmic fairness solutions in health care machine learning. Lancet Digital Health https://doi.org/10.1016/S2589-7500(20)30065-0

[138] Giovanola B & Tiribelli S (2022) Beyond bias and discrimination: redefining the AI ethics principle of fairness in healthcare machine-learning algorithms. AI and Society https://doi.org/10.1007/s00146-022-01455-6

alongside measures that both respect for personhood and promote equality of opportunity (e.g. policies for widening access, inclusive and accessible communications about AI).

3.1.2. Fairness and equity: relevance to TVB-Cloud

TVB-Cloud performed secondary research using data generated through clinical research on neurodegenerative diseases including Alzheimer's disease. However, the project did not involve any new clinical research studies. On the other hand, TVB-Cloud involved extensive processing of personal data, as well as the development, deployment and use of AI systems. In particular, TVB-Cloud developed a platform and processes for personalised brain simulations, currently useable by scientists and clinicians for the purposes of research and innovation – and in the future, to enable personalised decision-making, such as predicting risk, determining a diagnosis, or identifying potential treatments. Fairness and equity are therefore important considerations for TVB-Cloud at several levels.

Firstly, from the perspective of research users of TVB-Cloud, it is important to consider fair and equitable access to the results, outputs and innovations developed by the project. Open Access to publications is a cornerstone of the Horizon 2020 Model Grant Agreement, and as such TVB-Cloud publications are all clearly listed on the project website and are widely accessible online (https://virtualbraincloud-2020.eu/tvb-cloud-publications.html#publications-journals). A major innovation developed by TVB-Cloud is its Virtual Research Environment (VRE), which provides permissioned access to bona fide researchers, allowing them to find, securely access and share data. The VRE provides a wide range of support services and compute power for users, acting as an enabler for research which may be of particular benefit for scientists with fewer resources – thus enhancing fairness and equity.

The multi-scale brain simulations developed by TVB-Cloud are constructed using AI and as such algorithmic fairness is a relevant concern for the project. Algorithmic bias is an acknowledged limitation of projects like TVB-Cloud which are reliant on secondary analyses and uses of clinical research data; in the neurodegeneration research field, a recent empirical analysis demonstrated how inclusion of demographic features in a machine learning model to predict conversion to dementia resulted in a lowering of positive predictive values for Hispanic and non-white individuals (Sahin et al, 2022[139]). This has clear implications for the clinical translation of these models, and emphasises the importance of addressing underrepresentation in neurodegenerative disease research studies, which could benefit future iterations of TVB-Cloud brain simulations (Wang et al, 2022[140]). Algorithmic bias is dealt with in

[139] Sahin D, Jessen F, Kambeitz J, ADNI (2022) Algorithmic fairness in biomarker-based machine learning models to predict Alzheimer's dementia in individuals with mild cognitive impairment. Alzheimers & Dementia. https://doi.org/10.1002/alz.062125

[140] Wang X, Zhang R and Zhu R (2022) A brief review on algorithmic fairness. Management System Engineering. https://doi.org/10.1007/s44176-022-00006-z

greater detail in the section on non-discrimination below. Moreover, while TVB-Cloud innovations are not currently available for use by clinicians or patients, there are particular concerns linked to fairness and equity when considering deployment of e.g. decision support tools that will be developed using the brain simulations. As emphasised by Giovanola & Tiribelli (2022), respect of personhood and equity are important components of algorithmic fairness, and future development of TVB-Cloud innovations should incorporate measures which take account of structural inequalities and increase access, awareness and understanding for both clinicians and patients.

3.2. Informed consent

In D2.1, we explained that informed consent is essential to maintain legal and ethical compliance for health research participants. Informed consent is codified in the Declaration of Helsinki, which states "*Each potential (research) subject must be adequately informed of the aims, methods, sources of funding, any possible conflicts of interest, institutional affiliations of the researcher, the anticipated benefits and potential risks of the study and the discomfort it may entail, post-study provisions and any other relevant aspects of the study. The potential subject must be informed of the right to refuse to participate in the study or to withdraw consent to participate at anytime without reprisal. Special attention should be given to the specific information needs of individual potential subjects as well as to the methods used to deliver the information."* There is a legal dimension to informed consent in the healthcare setting: if a doctor does not inform a patient of risks they know are important and material to the decision at hand, they may be legally liable (Tawose, 2008[141]). Consent is also one of the legal grounds for processing personal data under the GDPR (as addressed in previous sections), and is a fundamental condition under which an individual can be included in a clinical trial in the Clinical Trials Regulation 536/2014 (CTR).

Informed consent allows individuals to control whether or not they participate in clinical research and, secondly, that they can choose to do so when the research is consistent with their values, interests and preferences. Informed consent meets the core ethical requirement of respect for autonomy, by enabling individuals to exercise their rights to self-government and self-determination. Indeed, the first question on the ethics issue checklist for H2020 applicants relates to whether informed consent has been obtained; applicants are asked to provide copies of informed consent forms and supporting study documentation. In addition, they must provide details on the informed consent procedures and, where vulnerable individuals or groups are involved, they must "*demonstrate appropriate efforts to ensure fully informed understanding of the implications of participation*". This references the concept of decisional

[141] Tawose MO (2008) The legal boundaries of informed consent. *Virtual Mentor.* doi: 10.1001/virtualmentor.2008.10.8.hlaw1-0808.

capacity, which refers to the ability of a potential research subject to understand and logically process the information that is necessary to make an informed decision regarding study participation. Research decisional capacity is not dictated by a person's diagnosis, socioeconomic or other profile; is context-specific; and may fluctuate over time. For example, a person may have capacity to consent to a low-risk research protocol, but may not be capable of understanding and processing information for a high-risk or complex protocol.

In the context of neurodegenerative disease research, one of the most important determinants of capacity is cognition. Any condition that affects a person's cognitive abilities may impair (or alter) that person's decision-making capacity. Providing informed consent requires a degree of reasoning, the ability to assimilate and evaluate information, and understand the risks and benefits of research. Many people with neurodegenerative diseases may, at some point, lack the capacity to reason, deliberate, and rationally understand, due to progressive cognitive impairment. This is an important consideration for projects such as TVB_Cloud, which focus on neurodegenerative diseases that are associated with cognitive impairment. However, assuming a lack of capacity solely because of a neurodegenerative disease diagnosis (or other characteristics) would be discriminatory, going against the fundamental ethical principles of justice, solidarity, and respect for autonomy. Of note, many people with dementia retain the capacity to express their desire to participate in research, even when they may have lost the capacity to make decisions about their financial affairs, for example. Even when people with dementia have lost legal capacity to provide informed consent, it is still possible to understand their preferences and interests through supported decision-making. The concept of supported decision making is about providing the necessary support for someone to make a decision whereby that person retains their legal capacity, even though they would not have been considered capable of deciding in the absence of that support[142]. The person or people providing the support are not necessarily relatives but could be anyone whom the person trusts. This support might, for example, involve providing information, explaining issues, describing different possible consequences of various options or helping the person to communicate the decision. In practice, support options can include simplifying consent forms, providing visual or memory aids, using digital, dynamic consent interfaces, or involving family caregivers in facilitating explanations (Thorogood et al, 2018[143]). In addition, informed consent can be organised as an ongoing, iterative process involving techniques to enhance individuals' reasoning and understanding,

[142] Alzheimer Europe (2020) Legal capacity and decision-making: the ethical implications of lack of legal capacity on the lives of people with dementia.

[143] Thorogood A, Petaja-Leinonen AM, Brodaty H, Dalpe G, Gastmans C, Gauthier S, Gove D, Harding R, Knoppers BM, Rossor M, Bobrow M, on behalf of the Ageing and Dementia Task Team of GA4GH (2018) Consent recommendations for research and international data sharing involving persons with dementia. Alzheimer's & Dementia https://doi.org/10.1016/j.jalz.2018.05.011

rather than a one-off interaction and signature (Gove et al, Alzheimer Europe, 2019[144]). By using reasonable accommodations such as these, the capacity of people with dementia (and/or other disabilities) to provide informed consent, and their right to participate in research, can be maximised.

Measures such as proxy decision making can extend a person's autonomy into the future, beyond the point at which they lose the capacity for informed consent. For example, a 2011 JAMA Psychiatry study showed that a substantial proportion of people with AD who did not have the capacity to consent to clinical studies had preserved capacity to appoint a research proxy (Kim et al, 2011[145]). However, there may be divergences in opinion, agendas or a lack of transparency between participants and their research proxies, which can have a negative impact on the person's right to autonomy and self-determination. Moreover, an over-reliance on research proxies can propagate paternalistic attitudes, neglecting the possibility for people with cognitive impairment to provide assent for research participation; proxies may be more cautious, and studies have shown that caregivers (who often take on the role of research proxy) can underreport health-related quality of life measures compared to patient self-assessment[146]. Finally, the burden of being a research proxy should also be considered; being asked to take decisions and assess risks or benefits on behalf of a loved one could expose the proxy to psychological harm if, for example, the person with cognitive impairment experienced severe side effects due to research participation.

Advance research directives can also provide a way for people with diminishing capacity to ensure their voices are heard during decision-making processes in clinical research studies. Advance directives, which are legal documents, can be used to appoint a proxy, although it may be advantageous for research participants to have both an advance directive and a proxy who has the power to make decisions not covered in the directive. Some of the ethical arguments supporting the use of advance directives for persons with cognitive impairments and/or dementia are summarised in the 2005 Alzheimer Europe Position Paper on the Use of Advance Directives (Alzheimer Europe, 2005[147]) and addressed in the Alzheimer Europe's 2019 report[148] entitled "Overcoming ethical challenges affecting

[144] Alzheimer Europe (2019) Overcoming ethical challenges affecting the involvement of people with dementia in research: recognising diversity and promoting inclusive research. https://www.alzheimer-europe.org/sites/default/files/2022-01/05706%20Alzheimer%20Europe%20ethics%20report%202019_92.pdf

[145] Kim SYH, Karlawish JH, Kim M, Wall IF, Bozoki AC and Appelbaum P (2011) Preservation of the capacity to appoint a proxy decision-maker: implications for dementia research. JAMA Psychiatry. doi: 10.1001/archgenpsychiatry.2010.191.

[146] Arons AM, Krabbe PF, Scholzel-Dorenbos CJ, van der Wilt GJ and Olde Rikkert GM (2013) Quality of life in dementia: a study on proxy bias. BMC Medical Research Methodology https://doi.org/10.1186/1471-2288-13-110

[147] Alzheimer Europe (2005) Advance Directives: a position paper https://www.alzheimer-europe.org/sites/default/files/2021-10/Advance%20Directives%20-%20Position%20Paper%202005.pdf

[148] Alzheimer Europe (2019) Overcoming ethical challenges affecting the involvement of people with dementia in research: recognising diversity and promoting inclusive research. https://www.alzheimer-europe.org/resources/publications/2019-alzheimer-europe-report-overcoming-ethical-challenges-affecting

the involvement of people with dementia in research: recognising diversity and promoting inclusive research". Advance directives are an effective means of preserving the autonomy of people with dementia, allowing them to exercise their right to self-determination. However, there should be safeguards for the use of advance directives for participation in research, for example through the inclusion of details on the level of risk/benefit that would be acceptable; monitoring of wellbeing throughout the research study; and involvement of ethics committees with expertise in dementia issues.

3.2.1. Informed consent: relevance to TVB-Cloud

TVB_Cloud has performed extensive secondary research on data shared from neurodegenerative disease research studies. For example, in a study that was recently published in Cerebral Cortex (Petkoski et al, 2023[149]). TVB_Cloud researchers analysed anatomical and diffusion-weighted brain scans that were initially acquired for a 2015 study that developed an image processing pipeline to construct individualised virtual brains. A 2022 study published in Alzheimer's & Dementia used PET and MRI imaging data from 33 participants in the US-based Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort (Triebkorn et al, 2022[150]). Informed consent is therefore an important ethical consideration for the project.

The informed consent documentation used for ADNI participants exemplifies a form of consent called "broad consent". Broad consent is an alternative to study-specific consent, in which participants consent to a broad framework for future use of their data and samples. Broad consent requires many of the same elements as study-specific consent, such as a description of reasonably-foreseen risks and benefits; a statement explaining how confidentiality will be assured; and a statement explaining that participation is voluntary, with the option to withdraw or discontinue at any time (Maloy & Bass, 2020[151]). However, broad consent forms also include statements explaining that data and samples will be shared and re-used for research beyond the initial, primary use. For example, in the ADNI consent form, participants are informed that their data will be "*stored indefinitely (at the Laboratory of Neuroimaging at the University of Southern California) and shared for future research*." And that "*All of the [de-identified] research data will be made available to qualified investigators at other scientific*

[149] Petkoski S, Ritter P and Jirsa V (2023) White-matter degradation and dynamical compensation support age-related functional alterations in human brain. Cerebral Cortex https://doi.org/10.1093/cercor/bhac500

[150] Triebkorn P, Stefanovski L, Dhindsa K, Diaz-Cortes MA, Bey P, Bulau K, Pai R, Spiegler A, Slodkin J, Jirsa V, McIntosh AR, Ritter P, for ADNI (2022) Brain simulation augments machine-learning-based classification of dementia. Alzheimers & Dementia TRCI. https://doi.org/10.1002/trc2.12303

[151] Maloy JW and Bass PF (2020) Understanding broad consent. Ochsner J. DOI: 10.31486/toj.19.0088

institutions around the world for research purposes." (ADNI). Broad consent has the advantage of granting researchers permission to use data for a range of different research studies, which are not specified at the time of collection (Thorogood et al, 2018). On the one hand, this increases the potential for societal benefit by driving new research that maximises the use and utility of data – which could be viewed as an ethical imperative. On the other hand, broad consent raises concerns about privacy, thus potentially exposing the research participant to loss of confidentiality and harm that could arise from the sharing and re-use of data. Opponents of broad consent also argue that consent cannot be truly informed if the participant cannot know how their data will be used in the future; and that study-specific consent is the only way to fully respect the right to autonomy. However, others argue that broad consent could be particularly well-suited to research projects involving people with cognitive impairment, who may not be able to provide study-specific consent at a later date, but who would want their wishes regarding data sharing to be respected. Broad consent proponents also argue that study-specific consent is much more resource-intensive, requiring repeated requests to consent for new uses of data from research participants. As well as using resources that could be better used elsewhere, repeated requests may lead to consent fatigue; bias; and the exclusion of people who may no longer have decisional capacity, or who may be harder to contact due to socioeconomic or other factors.

As a project that has benefited from the secondary use of data, it is essential for TVB_Cloud to consider the aforementioned ethical implications of broad and study-specific consent. The TVB-Cloud partners contributing background data, in particular clinical research data, ensured that the informed consent allowed the use of this data for the TVB-Cloud project. More importantly, having created a platform that enables researchers to access, analyse and share clinical research data (the Virtual Research Environment/VRE), TVB_Cloud should consider how to leverage the benefits of broad consent, whilst respecting the ethical tenets that underpin study-specific consent. A recent analysis on broad consent in the context of international biobanking by Mikkelsen et al (2019[152]) proposes two additional elements to ensure that broad consent provides sufficient ethical protection for research participants. Firstly, they recommend a continuous ethical review process, which should assess each new study proposal to determine whether it falls within the boundaries specified in the initial broad consent documentation. This is an approach that is used fairly widely in data sharing initiatives: for example, to access ADNI data and samples, researchers must sign a data use agreement and complete an application form, which is reviewed by their Data Sharing and Publications Committee. Applications for sample use are reviewed by the ADNI Resource Allocation Review Committee and the Biospecimen Review Committee, which have specific policies and protocols to govern approvals. However, the creation of ethical oversight

[152] Mikkelsen RB, Gjerris M, Waldemar G & Sandoe P (2019) Broad consent for biobanks is best – provided it is also deep. BMC Medical Ethics https://doi.org/10.1186/s12910-019-0414-6

committees governing the re-use of data or samples is likely beyond the remit of the VRE, which has been designed as an enabling environment for data sharing, and does not assume controllership of data that is processed, shared or analysed within the platform.

A second recommended element is continuous communication with research participants, where possible. Mikkelsen states that this has a number of advantages, beyond supporting ethical compliance: "*First, it allows [the biobank] to build trust with participants. There is evidence that such trust is central to willingness to participate. Second, continuous information allows participants to keep abreast of the evaluations made in the ethical review process. This ensures that they are aware of the research that is taking place and allows them to evaluate whether their values continue to align with the wider activities of the biobank and consider whether they still wish to be enrolled*." While the VRE is not subject to the same degree of ethical compliance as biobanks (the subject of Mikkelsen's recommendations), it may be feasible for the VRE to develop communications that disseminate some of the research activities and outputs from the platform, in collaboration with the researchers involved. This could have a number of benefits: firstly, researchers would be able to feed back information on how data has been used to the participant community and, secondly, dissemination could help build public trust in data sharing. Qualitative research studies show that despite broad agreement on the societal value of data sharing, patients and participants still have concerns about the potential loss of privacy - and a perceived lack of transparency in how and when data is shared (Alzheimer Europe report on Data Sharing in Dementia Research, 2021[153]). The value of education and awareness to foster trust and transparency also features prominently in ethical frameworks for AI. For example, the Ethics Guidelines for Trustworthy AI developed by the EC HLEG states: *"Trustworthy AI encourages the informed participation of all stakeholders. Communication, education and training play an important role, both to ensure that knowledge of the potential impact of AI systems is widespread, and to make people aware that they can participate in shaping the societal development."* This analysis makes it clear that communication and dissemination of VRE outputs and activities could be a valuable way to address ethical challenges. Trust and transparency will be addressed in more detail in section 3.5 and 3.6 below.

An additional ethical challenge with relevance for TVB_Cloud is how to deal with consent for secondary use of data in situations where capacity is lost or fluctuates – as a result of cognitive impairment, for example. There are a number of guidelines that can be followed by researchers and research initiatives. National Health Research Authorities (HRA) in certain countries have made recommendations to support the inclusion of participants who may lose capacity during participation in clinical trials of investigational medicinal products. The British National Health Service HRA states that consent

[153] Alzheimer Europe (2021) Data sharing in dementia research: the EU Landscape https://www.alzheimer-europe.org/policy/positions/data-sharing-dementia-research-eu-landscape

discussions should pro-actively address loss of capacity in situations where there is a significant risk of this arising. However, they also note that consent to participate in a study is presumed to remain legally valid after loss of capacity, provided the protocol does not change significantly. Nevertheless, researchers have a legal and ethical obligation to consider possible benefits and harms of continued participation given the participants' current situation, and seek approval from an appropriate research ethics committee as necessary. It is also important that researchers remain alert to, and respect, signs of not wishing to continue, whilst taking care not to be overly protective or paternalistic, which risks depriving people with the possibility of continued participation when this is aligned with their desires. Broad consent mechanisms, as discussed earlier, can obviate this need, supporting continued participation in research – and sharing of derived data, via platforms such as the VRE. To alleviate any remaining ethical concerns with broad consent mechanisms and loss of capacity, ELSI specialists have recommended that consent forms include a provision stating that consent to research or data sharing should be respected after a loss of capacity, unless the participant dissents or withdraws of their own accord. In addition to broad consent, people who have lost capacity can also be supported to make their own decisions about research participation and data sharing by using research proxies, or by referring to precedent consent through an advanced research directive (as discussed in the previous section). These options should be considered in the event that TVB_Cloud performs prospective research studies in the future; guidance could also be provided to users of the VRE.

3.3. Non-discrimination

In D2.1, non-discrimination was identified as one of the key ethical concerns for TVB_Cloud. Non-discrimination is one of the fundamental values of the EU, and therefore is enshrined in foundational EU instruments. However, there are currently only two EU legal instruments applicable to non-discrimination in the healthcare sector and they are limited to two types of discrimination: gender and race. Therefore, ethical codes and frameworks are particularly important to ensure the respect of the fundamental rights of participants in health research.

With specific reference to clinical research studies contributing datasets to the TVB-Cloud project, it should be recognised that people living with neurodegenerative disease and their caregivers have the right to be free from discrimination based on any grounds such as age, disability, gender, race, sexual orientation, religious beliefs, health status and also directly because of their health conditions. This extends to the right to participate in clinical research. However, a systematic review of clinical trials on disease-modifying therapies for Alzheimer's disease showed that participants were predominantly White (median: 94%), with exclusion criteria such as psychiatric or cardiovascular disease, and

requirements such as obligatory caregiver attendance (Franzen et al, 2021[154]). These criteria can lead to a systemic underrepresentation of people who are older, female, ethnically diverse, less educated, and less wealthy, aggravated by recruitment and retention strategies that do not account for attitudinal factors such as mistrust of research/health professionals, as well as differing levels of health literacy. A recent publication from the Clinical Trials in AD Taskforce highlighted that longitudinal, observational studies also suffer from a lack of diversity and inclusion, with a high proportion of white and highly-educated participants (Raman et al, 2022[155]). For example, the US-based ADNI study, which provides its data without embargo to all bona-fide scientists across the world, is cited in almost 4,000 research publications, and has been extensively leveraged by neurodegeneration researchers, data scientists and AI developers. However, the current composition of the ADNI participant group is 89% White, 8% Hispanic, 5% Black/African-American, and 3% Asian, with 85% of participants having an education level of 12 years or higher – while the general US population is 57.8% White, 18.7% Hispanic/Latino, and 12.1% Black/African-American. Lack of diversity is therefore a particular concern for projects such as TVB_Cloud, which rely on shared data from clinical studies to construct brain simulations, develop and validate decision support tools, and create disease progression models. Careful consideration of data diversity, and methods to mitigate algorithmic bias, would benefit the future development of TVB-Cloud brain simulations.

Connected to the acknowledged lack of diversity discussed in the previous paragraph, it is important to consider the risk of bias. A recent systematic review evaluated 92 studies using interpretable machine learning in dementia research, revealed that 67 studies employed the ADNI dataset – which has many features that make it very attractive for machine learning research (Martin et al, 2023[156]). However, they explained that this limits the generalisability of machine learning algorithms, particularly when they are applied in more diverse hospital, memory-clinic or community-based settings. This highlights the very real risk of bias in medical AI for neurodegenerative disorders such as dementia. Compounding the bias arising from a lack of diversity in clinical research studies (linked to inequitable recruitment and retention strategies, among other causes) is bias linked to the use of certain clinical tests. Widely-cited examples include racial bias in pulse oximetry sensors, and ethnic non-equivalence in performance on

[154] Franzen S, Smith JE, van den Berg E, Rivera Mindt M, van Bruchem-Visser RL, Abner EL, Schneider LS, Prins ND, Babulal GM and Papma J (2021) Diversity in Alzheimer's disease drug trials: the importance of eligibility criteria. Alzheimers & Dementia. 10.1002/alz.12433

[155] Raman R, Aisen PS, Carillo MC, Detke M, Grill JD, Okonkwo OC, Rivera-Mindt M, Sabbagh M, Vellas B, Weiner M & Sperling R (2022) Tackling a major deficiency of diversity in AD therapeutic trials: a CTAD Task Force report. J. Prev. Alzheimers Dis. 10.14283/jpad.2022.50

[156] Martin SA, Townend FJ, Barkhof F and Cole JH (2023) Interpretable machine learning for dementia: a systematic review. Alzheimer's & Dementia. https://doi.org/10.1002/alz.12948

the mini-mental state exam (MMSE) (Ng et al, 2007[157]). Natural Language Processing algorithms to identify language impairment as an early sign of AD were revealed to be biased towards Canadian English, exposing French-Canadian individuals or people from other language backgrounds to the risk of discrimination. Finally, bias may also be present in the access, deployment and implementation of medical AI. For example, even within the EU there is unequal access to the compute power that drives AI, and a 2022 WHO Europe scoping review revealed inequities in use and access to digital health technologies between urban and rural areas; younger and older individuals, and white, English-speaking Europeans compared to those from an ethnic minority with language barriers (WHO Europe report on equity within digital health technology in the WHO European Region, 2022[158]). Introducing AI into healthcare without considering and addressing these systemic inequities runs the risk of widening existing healthcare disparities. Mitigating each step of this bias cascade is crucial to making more accurate, sensitive and equitable AI-based tools that can benefit all groups within society.

Unsurprisingly, non-discrimination and avoidance of bias feature prominently in ethical recommendations and guidelines for AI. The 2021 UNESCO Recommendations on the Ethics of Artificial Intelligence[159] identifies "ensuring diversity and inclusiveness" as one of its core values, stating that: "*Respect, protection and promotion of diversity and inclusiveness should be ensured throughout the life cycle of AI systems, consistent with international law, including human rights law. This may be done by promoting active participation of all individuals or groups regardless of race, colour, descent, gender, age, language, religion, political opinion, national origin, ethnic origin, social origin, economic or social condition of birth, or disability and any other grounds.*" Non-discrimination, which is linked to fairness, is identified as one of its core principles*: "…the benefits of AI technologies [should be] available and accessible to all, taking into consideration the specific needs of different age groups, cultural systems, different language groups, persons with disabilities, girls and women, and disadvantaged, marginalized and vulnerable people or people in vulnerable situations."* The Recommendations also link non-discrimination to algorithmic equality, stating that: "*AI actors should make all reasonable efforts to minimize and avoid reinforcing or perpetuating discriminatory or biased applications and outcomes throughout the life cycle of the AI system to ensure fairness of such systems. Effective remedy should be available against discrimination and biased algorithmic determination."* Similarly, non-discrimination and avoidance of bias is emphasised in the EU Ethics Guidelines for Trustworthy Artificial Intelligence. Equality, non-discrimination and solidarity are identified as a family of fundamental rights that are

[157] Ng PT, Niti M, Chiam PC and Kua EH (2007) Ethnic and educational differences in cognitive test performance on mini-mental state examination in Asians. Am. J. Geriatr. Psychiatry. 10.1097/01.JGP.0000235710.17450.9a
[158] World Health Organisation – Europe (2022) Equity within digital health technology within the WHO European Region: a scoping review. https://www.who.int/europe/publications/i/item/WHO-EURO-2022-6810-46576-67595
[159] UNESCO (2021) Recommendation on the ethics of AI. https://unesdoc.unesco.org/ark:/48223/pf0000380455

"particularly apt to cover AI systems", with the guidelines stating that: "*In an AI context, equality entails that the system's operations cannot generate unfairly biased outputs (e.g. the data used to train AI systems should be as inclusive as possible, representing different population groups). This also requires adequate respect for potentially vulnerable persons and groups, such as workers, women, persons with disabilities, ethnic minorities, children, consumers or others at risk of exclusion.*" In detailing the principle of fairness, the guidelines explain that there is a substantive dimension to fairness linked to non-discrimination, which implies *"a commitment to ensuring equal and just distribution of both benefits and costs, and ensuring that individuals and groups are free from unfair bias, discrimination and stigmatisation*."

3.3.1. Non-discrimination: relevance to TVB-Cloud

Although TVB_Cloud did not initiate any new clinical research studies, it did involve the re-use of data from studies on neurodegenerative diseases including Alzheimer's disease, and the development, deployment and use of AI systems to create personalised brain simulations. These multi-scale brain simulations are constructed using AI and, as indicated in the previous section on fairness, algorithmic bias is therefore a relevant concern for the project. Algorithmic bias is an acknowledged limitation of projects like TVB-Cloud which are reliant on secondary analyses and uses of clinical research. However, there are tools and processes that can help identify and address algorithmic bias and other sources of discrimination. For example, the Assessment List for Trustworthy Artificial Intelligence ALTAI)[160] was launched in July 2020 by the HLEG on AI, with a web-based tool to support the development of trustworthy AI in compliance with the ethical principles and recommendations laid out in their guidelines. Building on these guidelines, the ALTAI poses a series of questions to assess whether adequate measures have been taken to avoid unfair bias. For example:

- *Did you establish a strategy or a set of procedures to avoid creating or reinforcing unfair bias in the AI system, both regarding the use of input data as well as for the algorithm design?*
- *Did you consider diversity and representativeness of end-users and/or subjects in the data?*
- *Did you put in place educational and awareness initiatives to help AI designers and AI developers be more aware of the possible bias they can inject in designing and developing the AI system?*
- *Did you consult with the impacted communities about the correct definition of fairness, i.e. representatives of elderly persons or persons with disabilities?*

[160] EU HLEG: Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment (2020) https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment

A further tool to support algorithmic impact assessments was launched by the Finnish Prime Minister's Office in 2022, creating an assessment framework that combines the evaluation of discriminatory risks of AI systems with the promotion of equality (Finnish Assessment Framework for Non-discriminatory AI Systems, 2022[161]). The framework, which follows a lifecycle model (design – development – deployment) is built on a national mapping of AI applications in public use and an in-depth analysis of the discriminatory risks, taking into account legal frameworks such as the Finnish Non-Discrimination Act.

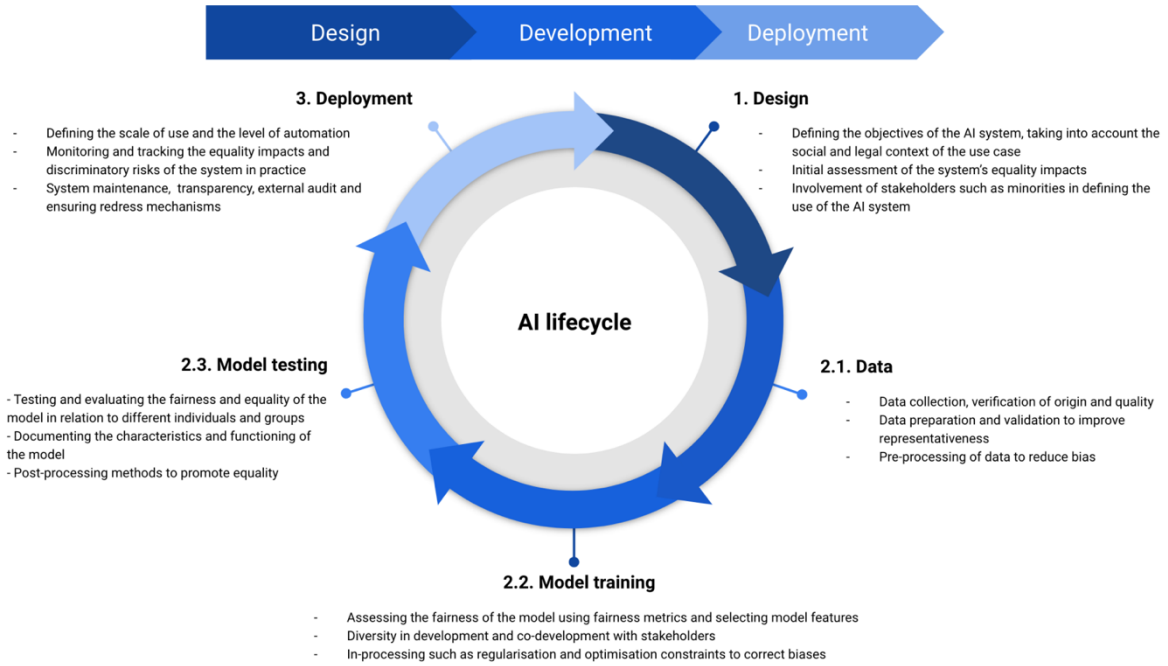


Figure 3.3.1. Overview of the Finnish assessment framework for non-discriminatory AI

These and other tools (such as metrics for measuring AI risks) are openly available to researchers, developers and other AI actors, and are listed in the OECD Catalogue of Tools & Metrics for Trustworthy AI[162], which provides a one-stop shop for enabling processes, mechanisms and practices. Consideration of, and use of these resources may help reduce the risk of discrimination and enhance fairness as TVB_Cloud undergoes further development.

Transparency features prominently in ethical frameworks for AI, and can also help identify biases and mitigate issues of fairness and discrimination. Algorithmic transparency is the principle that factors that influence the decisions made by AI algorithms should be visible, or transparent, to people who use,

[161] Finnish Government (in collaboration with Demos Helsinki, University of Turku and University of Tampere) An assessment framework for non-discriminatory AI (2022) https://demoshelsinki.fi/julkaisut/an-assessment-framework-for-non-discriminatory-ai/

[162] OECD Catalogue of tools & metrics for trustworthy AI: https://oecd.ai/en/catalogue/tools

regulate and are affected by these algorithms. Transparent algorithms can enhance non-discrimination by exposing possible biases in training data, by showing how AI learns over time, its boundaries and safety measures. On the other hand, technical developers have historically argued that complexity is inherent in highly accurate AI tools that are more flexible (e.g. black box deep learning models), and that there is a tradeoff between accuracy and explainability. Interestingly, studies indicate that the general public favours accuracy over explainability for AI used in the healthcare setting, but have a greater preference for explainability in non-healthcare contexts[163]. Underlining the importance of algorithmic transparency, whilst acknowledging the technical and implementation challenges that developers face, the European Commission launched the European Centre for Algorithmic Transparency (ECAT) in April 2023 as a branch of the Joint Research Centre. The ECAT, once fully operational, will provide scientific and technical expertise to support the supervisory and enforcement role for very large online platforms and search engines. While this may fall beyond the remit of TVB_Cloud, future development of TVB_Cloud tools will likely be subject to the AI Act, a flagship piece of EU legislation that will regulate the use of AI systems in Europe (see section 2.3.1. for a detailed description of the legislation and an analysis of its application to TVB_Cloud). A draft negotiating mandate was adopted by the European Parliament's Internal Market and Civil Liberties committees in early May 2023, and will be discussed in plenary in June. One of the aims of the AI Act is to ensure that AI systems are non-discriminatory using a human rights- and risk-based approach to categorise technologies, and at the recent committee meetings MEPs substantially amended the list of prohibited "unacceptable risk" technologies to further limit the risk of discrimination, including (among others) remote biometric identification systems in public spaces, predictive policing systems, and biometric categorisation systems using sensitive characteristics (e.g. gender, ethnicity, citizenship status, religion). As outlined in section 2.3.1, TVB_Cloud may be legally bound by the AI Act; in which case, careful consideration and mitigation of discrimination risks will be essential for further development and deployment.

3.4. Confidentiality

Confidentiality is one of the core duties of medical practice and a key concern for clinical researchers, linked to two ethical principles for health research: respect for persons, and minimising the risk for harm (Beauchamp and Childress, 1979). Unlike medical care, participation in research must be voluntary; and as such, researchers could be viewed as having a greater ethical obligation to protect the confidentiality of participants. Consequently, clinical research must be carried out in a way that secures and protects the privacy of participants, minimising the risk of unauthorised disclosure or reidentification. This

[163] Van der Veer SN, Riste L, Cheraghi-Soni S, Phipps DL, Tully MP, Bozentko K, Atwood S, Hubbard A, Wiper C, Oswald M, and Peek N (2021) Trading off accuracy and explainability in AI decision-making: findings from 2 citizen's juries. Journal of the American Medical Informatics Association. https://doi.org/10.1093/jamia/ocab127

obligation is reflected in paragraph 23 of the Declaration of Helsinki, which states that: *"Every precaution must be taken to protect the privacy of research subjects and the confidentiality of their personal information and to minimize the impact of the study on their physical, mental and social integrity*" - a concept which is also echoed by Article 56 on recording, processing, handling and storage of information of the Clinical Trial Regulation.

While confidentiality is an ethical duty for researchers and medical practitioners, privacy is a fundamental human right enshrined in the EU Charter of Fundamental Rights and the Treaty on the Functioning of the European Union, giving effect to individuals' right to privacy by providing them with control over the way information about them is collected and used. In Europe, data protection is regulated via the GDPR. The legal framework for TVB_Cloud is primarily focused on GDPR compliance, which is extensively discussed in the previous sections of this report. In this section of the framework, we will address the **ethical challenges** linked to confidentiality: in particular the risks of harm for participants in neurodegeneration research studies that could arise as a result of loss of confidentiality.

As indicated above, research participants have a right to confidentiality with regards to their personal health information, and a right to privacy. As well as psychological and physical harms, loss of confidentiality and privacy exposes research participants to a risk of stigmatisation, through negative stereotypes, prejudice and discrimination. Stigma is a recurring topic in the ethical discourse around neurodegenerative diseases such as Alzheimer's and dementia (Alzheimer Europe report on the Ethics of Dementia Research, 2011[164]). Stigma is a complex social phenomenon involving a process and being characterised by a set of components which are constantly perpetuated within society. This causes a person's sense of self and social status to be "tainted" and/or devalued[165], for example when linked with a disease that has negative health connotations. Diseases come to have negative connotations because of the meanings that are associated with the conditions (i.e. the socially salient attribute) however, these connotations are not necessarily or primarily health related, and may be linked to prejudice and negative stereotypes (such as obese people being lazy, people with mental disorders being dangerous, people with dementia symbolising a loss of self). Public stigma involves the identification and labelling of people with a socially salient attribute (such as having dementia), attaching negative stereotypes, considering people with that attribute as being 'other' (i.e. not like 'us'), devaluing them and discriminating against them (including denial of equal opportunities as well as social distancing) (Link and Phelan 2001[166]). The attribute is not inherently stigmatising but comes to be considered as such

[164] Alzheimer Europe (2011) The ethics of dementia research. https://www.alzheimer-europe.org/sites/default/files/alzheimer_europe_ethics_report_2011.pdf
[165] Goffman E (1963) Stigma: notes on the management of spoiled identity. Eds. Penguin Books
[166] Link BG and Phelan JC (2001) Conceptualising Stigma. Annual review of Sociology https://doi.org/10.1146/annurev.soc.27.1.363

because of the meanings attached to it which are socially salient and generally involve a perception of some kind of threat (e.g. to individuals, society or morality) (Stangor and Crandall 2000[167]). Self-stigma involves the internalisation of negative stereotypes, resulting in low self-esteem and the anticipation of discrimination. Internalised stigma can lead to depression, avoidant coping and social withdrawal for people with dementia; while public or cultural sigma is linked to overt or subtle discrimination, exclusion, or patronising attitudes towards people affected by AD and dementia. Stigma can also affect families or caregivers of people with AD and dementia, who may experience stigma by association with their loved ones.

Confidentiality risks are also highlighted in the 2019 EU Ethics Guidelines for Trustworthy AI, which identifies privacy and data protection as one of seven key requirements for trustworthy AI. According to these guidelines, AI systems must guarantee privacy and data protection throughout the data lifecycle, from the point that data is provided to the system by the user, to the eventual outputs of the AI systems for users. AI must also "*be lawful, complying with all applicable rules and regulations*"; the guidelines reference the GDPR, the European Convention on Human Rights, and the EU Charter of Fundamental rights, among others. To minimise confidentiality and privacy risks, the guidelines state that "*prevention of harm to privacy [also] necessitates adequate data governance that covers the quality and integrity of the data used, its relevance in light of the domain in which the AI systems will be deployed, its access protocols and the capability to process data in a manner that protects pri*vacy." To support implementation of the guidelines, an assessment list is incorporated in to the document, asking specific privacy and data protection questions to support the ethical implementation of AI-based innovations such as TVB_Cloud in compliance with the GDPR, for example:

- *Did you consider ways to develop the AI system or train the model without or with minimal use of potentially sensitive information or data?*
- *Did you take measures to enhance privacy, such as via encryption, anonymisation or aggregation?*
- *Did you establish oversight mechanisms for data collection, storage, processing and use?*

3.4.1.Confidentiality: relevance to TVB_Cloud

Loss of confidentiality or privacy is a major ethical concern for projects such as TVB_Cloud, which involve the secondary use of data from clinical research studies on neurodegenerative diseases. A further dimension to consider is the loss of confidentiality that could occur when data is uploaded, shared or analysed in the TVB_Cloud VRE. Loss of confidentiality could arise in connection with technical issues

[167] Stangor C and Crandall CS (2000) Threat and the social construction of stigma. In T. F. Heatherton, R. E. Kleck, M. R. Hebl, & J. G. Hull (Eds.), *The social psychology of stigma*

(e.g. security breaches), or due to reidentification of individuals, for example through re-association of identifiers with personal information, or through future linkage of individual genetic or brain imaging data with individuals. However, this risk must be accurately assessed, and balanced against the benefits of re-using clinical research data in the TVB_Cloud project - and of providing a data sharing platform that will enable research and innovation on neurodegenerative diseases. Indeed, it is widely acknowledged that the ethics of health research requires a context-sensitive approach, and a balanced risk assessment about likely harm to the individual (or his/her community) on the one hand, versus the right to perform research for societal benefit on the other hand. In its 2016 International Ethical Guidelines for Health-related Research[168], the Council for International Organisations of Medical Sciences (CIOMS) states:

"The researcher, sponsor and research ethics committee must ensure that risks to participants are minimised and appropriately balanced in relation to the prospect of potential individual benefit and the social and scientific value of the research."

Here, the term 'risk' encompasses physical, mental and emotional harms as well as concerns such as incurred costs and practical inconvenience – which could also be termed "burdens" of research participation. With reference to confidentiality, the risk of stigma is an important consideration for TVB_Cloud, as outlined in previous paragraphs. Estimating the risk of stigma is complicated by differing estimations and differing viewpoints; people have varying levels of subjective knowledge about neurodegenerative diseases such as dementia, and studies have shown that stigmatising attitudes also vary between individuals, demographic groups and societies, and over time (Werner and 2021[169]). Scott Kim and coauthors[170] highlight a further ethical challenge beyond the risk of stigma; the potential for direct psychological harm due to undesired disclosure of a disease diagnosis or disease risk status, particularly for conditions such as Alzheimer's disease, which have few or no effective treatment options. On the other hand, an overestimation of risk or harm by researchers may lead to the possible benefits of research being undervalued. Beyond the value of primary research, surveys show that data sharing and reuse is valued by the majority of research participants, linked to the potential for scientific advancement, greater research efficiency and improved health outcomes at the societal level (Alzheimer Europe report on Data Sharing in Dementia Research, 2021). In this view, it could be argued that data sharing and reuse, for example using systems such as the VRE, is a moral and ethical

[168] Council for International Organisations of Medical Sciences (CIOMS) in collaboration with the WHO (2016) International Ethical Guidelines for Health-related Research Involving Humans https://cioms.ch/wp-content/uploads/2017/01/WEB-CIOMS-EthicalGuidelines.pdf

[169] Werner P and Kim S (2021) A cross-national study of dementia stigma among the general public in Israel and Austria. J. Alzheimers. Dis. 10.3233/JAD-210277

[170] Kim SYH, Karlawish J and Berkman BE (2015) Ethics of genetic and biomarker test disclosures in neurodegenerative disease prevention trials. Neurology: https://doi.org/10.1212/WNL.0000000000001451

imperative. Consequently, projects such as TVB_Cloud must carefully evaluate the risks linked to loss of confidentiality, and the likelihood that loss of confidentiality will occur, then balance these against the benefits of research to individuals and to society.

To ensure that data is protected and to reduce the likelihood of loss of confidentiality, TVB_Cloud has adopted a number of technical and organisation measures, such as authentication and authorisation procedures for project- and role-based access to the Virtual Research Environment; a federated identity management system; and a segregated "Green Room" area to capture and pre-process sensitive data to make it analysis-ready. As well as reducing the risk of privacy breaches, the VRE actively supports researchers in mitigating the risk of reidentification. These systems and processes have been audited by an external agency, which has certified that the VRE is GDPR service ready. As a further layer of good practice, and to enhance transparency, future development of TVB_Cloud could consider what types of additional information could be provided to users to explain how data is used or shared in the VRE; and how to disseminate this information to research participants and data subjects.

3.5. Transparency

Transparency is central to the ethical practice of research. In clinical studies, the ethical requirement for transparency spans the full research pipeline, from clear documentation on the study protocol and consent conditions (see previous section), to disclosure and eventual sharing of the results following study completion. For example, research ethics requires that participants are provided with full transparency on the study purpose, procedures, requirements, risks and benefits, while clinical trial regulations oblige sponsors to register studies on publicly-accessible databases (EudraCT in the EU; Clinicaltrials.gov in the US). These transparency rules, primarily designed to meeting ethical and legal requirements for research, can also help to build public trust in research, avoid study duplication, enable research participation, foster innovation and collaboration, and inform decision-making, among other benefits. Acknowledging the importance of research transparency, many countries have created strategies and rules to make transparency the norm in research. For example, the Health Research Authority (HRA) reviews and regulates all health and social care research being carried out in the UK. Their research transparency strategy requires all clinical trials and most interventional research research studies to be publicly registered prior to launch; asks investigators to submit a final report within 12 months of the end of a study; and expects that clinical investigators will make efforts to inform study participants on the research findings.

As stated in several of the previous sections, transparency is also a core principle in many ethical frameworks and codes of conduct for AI. Transparency can help build trust in AI, promote adoption, clarify liability and, additionally, can help identify biases and mitigate issues of fairness and

discrimination. Unsurprisingly, a 2021 Opinion[171] from the Human Brain Project's Ethics and Society subproject identifies transparency as the most prevalent principle in these documents, defining transparency as "*the need to make decision-making processes accessible to users, so they can understand and judge how an autonomous system has reached a certain decision*." As a case in point, the HLEG Ethical Guidelines for Trustworthy AI state that transparency is a crucial component for achieving Trustworthy AI, encompassing three elements: traceability, explainability, and open communication. According to these guidelines, explicability is one of the four core ethical principles of AI systems, stating that: *"[AI] processes need to be transparent, the capabilities and purpose of AI systems openly communicated and – to the extent possible – explained to those directly and indirectly affected*." Acknowledging that this may be challenging for "black box" algorithms, the guidelines go on to state: "other explicability measures (e.g. traceability, auditability and transparent communication on system capabilities) may be required, provided that the system as a whole respects fundamental rights." Transparency and explainability also feature in EU regulations; for example, Recital 71 of the GDPR states: "*The data subject should have the right not to be subject to a decision…evaluating personal aspects relating to him or her which is based solely on automated processing and which produces legal effects concerning him or her or similarly significantly affects him or her, such as automatic refusal of an online credit application or e-recruiting practices without any human intervention (…) such processing should be subject to suitable safeguards, which should include specific information to the data subject and the right to obtain human intervention, to express his or her point of view, to obtain an explanation of the decision reached after such assessment and to challenge the decision."*

3.5.1. Transparency: relevance to TVB-Cloud

Explainable AI (XAI) is an emerging concept that is steadily achieving greater prominence in the healthcare field. XAI refers to a set of methods and processes that helps users understand and trust the outputs and results given by AI-based systems. XAI provides causal and logical explanations for the decisions made by AI algorithms; for example, explaining which features are assigned a greater weight in AI models. These explanations can be surfaced to different stakeholders interacting with the AI system – doctors using decision support tools, for example, or patients receiving the results of risk prediction algorithms (McDermid et al, 2021[172]). As such, XAI can enhance transparency, and help build trust. XAI also has the potential to enable stakeholders to exercise their autonomy: for example, appropriate explanations could support informed consent. However, different stakeholders may have different explanatory requirements, and varying levels of digital literacy; in the case of TVB-Cloud, stakeholders

[171] Human Brain Project, Ethics and Society subproject (2021) Opinion on Trust and Transparency in Artificial Intelligence. https://zenodo.org/record/4588648

[172] McDermid JA, Jia Y, Porter Z and Habli I (2021) Artificial intelligence explainability: the technical and ethical dimensions. Philos. Trans. A Math Phys. Eng. Sci. https://doi.org/10.1098/rsta.2020.0363

may have neurodegenerative disease, and could be affected by cognitive impairment. This interpersonal variation means that XAI risks of exacerbating inequalities, if person-centered methods aren't used to communicate with stakeholders. An additional ethical challenge to consider is the risk of unjustified trust in a XAI system if the explanations aren't sufficiently accurate, detailed or understandable. These are all important ethical considerations for TVB-Cloud and other projects developing AI-based systems and innovations.

A further ethical concept linked to transparency and explainability is the principle of accountability. In standard clinical care, and in clinical research, there is clear accountability for doctors and clinical researchers. For example, patients expect their doctors to exercise good judgment, act with competence, and make decisions that are in their best interest. However, accountability is less clear when it comes to AI-based systems, particularly when they are deployed in the context of clinical care or research. The OECD's value-based principles for AI defines accountability as "*the expectation that organisations or individuals will ensure the proper functioning, throughout their lifecycle, of the AI systems they design, develop, operate or deploy, in accordance with their roles and applicable regulatory frameworks*." However, it is challenging to assign moral accountability in the deployment of AI-based systems for decision support in healthcare – one of the potential future applications of TVB-Cloud. Questions arise around how far a clinician is accountable for patient harm that may arise as a result of a decision influenced by an AI-based system. While the clinician can ultimately choose whether or not to act on a recommendation from a decision support system, they may not fully understand how the system has reached the conclusion, and the choice to implement the recommendation may also be impacted by structural factors not related to the patient themselves, such as the amount of time a clinician has to devote to the patient (Habli et al, 2020[173]). These factors can also influence their willingness or resistance to use AI to support decision-making. On the other hand, the AI developer will not be present at the point of clinical decision-making – and unlike clinicians, which are held to account through regulatory frameworks and by professional bodies, there are currently no legal tools to assign accountability and apply punitive or corrective measures (Smith, 2021[174]). This highlights the challenge of determine moral accountability for decision support tools such as those envisaged for future TVB-Cloud developments, and underlines the need for healthcare systems, policymakers and regulators to develop processes and instruments to help assign accountability to different actors.

3.6. Trustworthiness, and stakeholder involvement in TVB-Cloud

[173] Habli I, Lawton T & Porter Z (2020) AI in healthcare: accountability and safety. Bull. WHO https://doi.org/10.2471%2FBLT.19.237487

[174] Smith H (2020) Clinical AI: opacity, accountability, responsibility and liability. AI & Society https://doi.org/10.1007/s00146-020-01019-6

Co-design has the potential to enhance explainability – and increase transparency - by involving end users in early stages of the development of AI solutions. The value of consulting and involving stakeholders (clinicians and patients, for example) is highlighted in the HLEG Ethics Guidelines, which state: "*In order to develop AI systems that are trustworthy, it is advisable to consult stakeholders who may directly or indirectly be affected by the system throughout its life cycle. It is beneficial to solicit regular feedback even after deployment and set up longer term mechanisms for stakeholder participation, for example by ensuring workers information, consultation and participation throughout the whole process of implementing AI systems."* Similarly, the UNESCO Recommendations on the Ethics of AI states: "*Participation of different stakeholders throughout the AI system life cycle is necessary for inclusive approaches to AI governance, enabling the benefits to be shared by all, and to contribute to sustainable* development (…) *Measures should be adopted to take into account shifts in technologies, the emergence of new groups of stakeholders, and to allow for meaningful participation by marginalized groups, communities and individuals and, where relevant, in the case of Indigenous Peoples, respect for the self-governance of their data*."

Public involvement in TVB-Cloud

TVB-Cloud involved end users in two different workstreams, meeting these ethical recommendations. The first workstream, which is reported in deliverables from WP6, involved several consultations with Alzheimer Europe's European Working Group of People with Dementia (EWGPWD), to understand the feasibility and acceptability of a rehabilitation gaming system (RGS) for people with cognitive impairment. The EWGPWD comprises 13-15 people from across Europe, who are living with different types of dementia. Members are nominated by national Alzheimer's Associations to serve for a 2 year period, supported by caregivers, friends or relatives. The group is moderated by Dr. Dianne Gove and Dr. Ana Diaz, and consultations are carried out in the framework of patient and public involvement (or PI). PI is about carrying out research and developing policies with, or by, members of the public and patients rather than on or for them as mere participants (Gove et al, 2018[175]). PI can promote the transparency, validity and legitimacy of clinical research projects, by integrating the perspectives, needs and values of the ultimate beneficiaries of health research: patients, and the public.

[175] Gove D, Diaz-Ponce A, Georges J, Moniz-Cook E, Mountain G, Chattat R, Oksnebjerg L, and the EWGPWD (2018) Alzheimer Europe's position on involving people with dementia in research through PPI (Patient and Public Involvement). Aging & Mental Health. https://doi.org/10.1080/13607863.2017.1317334

TVB-Cloud consultations with the European Working Group of People with Dementia

Briefly, the RGSapp consists of a smartphone-based interface that allows people to play AI-based "brain training" games that may help with rehabilitation of motor and cognitive skills, such as executive function, memory and attention. RGSapp has been developed and deployed in clinics for stroke patients, and was undergoing further development in TVB-Cloud for use by people with cognitive impairment. Through a series of PI consultations, the EWGPWD provided feedback on the initial 2D RGSapp interface, and a 3D, augmented reality interface. The consultations allowed the RGSapp developers to identify features that made the games harder or easier to play, and ways to increase motivation and engagement of users with the interfaces. For example, the updated RGSapp now includes video feedback and instruction, replacing text instructions which some found confusing. To provide feedback on results, which could incentivise users to use the app, the RGSapp developers have created a virtual coach to provide motivational messages.

TVB-Cloud also involved stakeholders in the development of the present deliverable, obtaining feedback on two key themes discussed in previous sections of the ethical framework: trust, and transparency, in the context of data sharing and AI. These topics have particular relevance to the VRE – as a platform for data sharing – and the AI-based brain simulations, which were developed with shared health data. In April 2023, TVB-Cloud organised a consultation with the Alzheimer Europe European Dementia Carers Working Group (EDCWG) in Brussels. The EDCWG is composed of 13 current or former caregivers, relatives or supporters for people with dementia, and like the EWGPWD, the EDCWG takes an active part in consultations in the context of public involvement for research projects, providing their unique insights to advise and improve research.

During the consultation, a series of questions on trust, trustworthiness and data sharing were addressed. For example:

- What does trust mean to you in relation to people (e.g. doctors, researchers) and machines (e.g. cars, computers), compared to AI?
- How does trust change when you are trusting on behalf of someone else?
- What factors can enable trust? Conversely, what factors break trust?

The main themes from responses are summarised below, divided into short sections that link with the ethical issues identified in the previous pages.

Trust in people

Members explained that there can be implicit trust in doctors, and that this is reinforced when there is clear expertise and competence. Contact with doctors is often face-to-face, which can also enhance trust; trust is a "human feeling", which is hard to build when things are communicated by phone, online, or in an impersonal manner. For researchers, their credentials (e.g. as an expert professor) and the reputation of their institution can confer a certain level of trust (e.g. Universities of Oxford or Cambridge, in the UK), but similarly face-to-face contact and spending time with people/research participants are more important factors for building trust. Peer recommendation can increase the likelihood of trusting

a researcher, or indeed a doctor; members provided the example of having a recommendation from a family member, a "trusted source" of information.

In line with many of the ethical frameworks discussed in this deliverable, transparency and clarity of purpose were identified as important factors for trust in both doctors and researchers; if researchers provide clear information about what data is being used, and why, we are more likely to trust them. Similarly, doctors should devote time to clear explanations, which can help in building a good relationship with their patients.

Trust in AI

Members explained that trust in machines is more straightforward than trust in people; with machines such as computers or cars, you can "*turn it on and see if it works*", and you can get a sense of reliability – which enhances trust. Machines are also created on a much larger scale, whereas people are unique. With machines (and potentially AI) there is also the possibility of brand recognition helping to build trust, for example with Tesla cars: it is a recognisable brand, and we can rationalise that if one works well, we can trust that the others will also function similarly.

Similar to people, trust in machines and AI was viewed as context-dependent. We are more likely to trust machines or AI that are providing a service - or are designed to contribute to the common good. However, trust can be negatively impacted by hacks, data leaks or similar scandals (e.g. hack of the Irish health data system), and this can impact your likeliness to trust other systems in the same category.

Trusting on behalf of someone else

Several members shared that they had higher expectations and demands when it came to trusting a person, machine or AI on behalf of the person they care for. There are higher standards for trust when you are responsible for someone else, and similarly there are greater feelings of guilt when that trust is broken, particularly when it impacts negatively on the person you care for. Members explained that they would have more questions and would need more reassurance that trust is merited, as there is more pressure when making decisions on behalf of someone else, for whom you have a duty of care. Members noted that there is an imbalance of power when caring for someone who has dementia, which is different to e.g. caring for a child, as the person with dementia was previously able to make their own informed decisions.

In the second part of the consultation, members provided feedback on barriers and enablers of trust in AI and datasharing through discussion and also on post-it notes. Generally, the feedback reinforced the idea that trust is context-dependent. As such, members felt that barriers and enablers for trust in data

sharing and AI would vary depending on the intended use or AI application. Uses that could benefit the individual, or have wider benefit to society, would be easier to trust, and may require fewer assurances of trustworthiness. On the other hand, commercial uses of health data were inherently less trustworthy, although there was an acknowledgement that pharmaceutical companies may be less interested in tracing individuals than e.g. facebook.

Confidentiality

Measures to protect confidentiality and to secure data (e.g. anonymisation; technical systems to protect from hacking or leaks) were identified as important enablers of trust in AI, as were certification processes which can verify the reliability of data sharing or AI systems. Members explained that they would be much more comfortable sharing their data if anonymity was guaranteed. As well as increasing data security, measures to protect shared data also show that researchers respect and value the contribution of research participants, which further increases trust.

This feedback reinforces our analysis in the previous sections, which identified confidentiality as one of the core ethical issues for TVB-Cloud. Confidentiality, and protection of privacy, is both an ethical and legal imperative, and measures to protect data and effectively prevent reidentification are crucial to ensure trust, particularly when the extent and efficacy of these measures are made clear through accessible communications to data subjects.

Transparency

Transparency was identified as an important enabler to trust, with particular reference to transparency about purpose (why the data is shared, how it will be used), reliability, and security measures. Human oversight was also identified as an enabler of trust; either in. ethical committees that approve research, through certification of AI, or via human interactions with data sharing or AI systems.

The perspectives of the EDCWG clearly highlight the value of transparency as an enabler for understanding, and vector for effective communications, to end users and beneficiaries of TVB-Cloud. The principle – and challenge – of accountability is also raised; how can we design systems where accountability is clearly established and understood by all parties? Transparency is defined by the Human Brain Project analysis as "*the need to make decision-making processes accessible to users, so they can understand and judge how an autonomous system has reached a certain decision*." This feedback identifies a role for ethical committees, clinical or other experts, and certification systems, in enhancing transparency and building in a layer of validation and accountability.

Informed consent

There were varying views on consent as an enabler of trust. Broad consent was viewed favourably if it is clear that participants can opt out, if the data remains anonymous, and if it is clear how the data is used. However, some members felt uncomfortable at the idea of broad consent, preferring specific consent and to "err on the side of caution" when agreeing to data sharing. Ethics committees (with representation of end users) were viewed as a "trusted intermediary" and policing system, particularly when research is being carried out by a trusted institution/research team.

The EDCWG perspectives on informed consent also draw on the ethical issues of transparency and accountability, and reinforce the importance of inclusive and clear communication to end users, data subjects, and beneficiaries of TVB-Cloud. While broad consent was viewed favourably by some, others were less positive. Balancing the application of broad consent with study-specific consent to research should consider the potential risks of research and sharing/reuse of data from research, in consultation with participants, patients and caregivers.

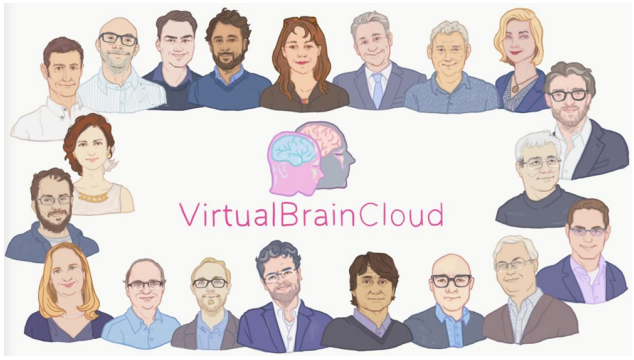
4. Description of work performed Task 2.5

Task 2.5 was targeted at participating in setting European best practices and industry standards in data protection of health and lifestyle related data. UNIVIE together with the support of TP21, AE, CHARITE, Fraunhofer, Eodyne were able to foster this goal by:

(1) organising a GDPR impact conference (Task 2.4) bringing together different stakeholders in the field to discuss data protection in health research with a special focus on the use of AI (see Deliverable D2.4)
(2) attending the EDPB Stakeholder Event on processing of personal data for scientific research purposes held on 30 April 2021, 10:00-16:00 (CET)
(3) presenting at a webinar organised by the KATY project (https://katy-project.eu/):

KATY-webinar 25.3.2021

universität wien



Lessons learned - so far

Presented by: Michael Cepic

25.03.2021 KATY-webinar

(4) presenting at the CPDP 2022 conference (https://www.cpdpconferences.org/) in the panel on "Regulating AI in Health Research and Innovation" with a presentation on "Preliminary considerations for the interplay between the propsals for an Artificial Intelligence Act and the European Health Data Space"

(5) organising workshops with Boehringer-Ingelheim, which themselves, are working on a code of conduct under the EFPIA framework (https://www.efpia.eu/news-events/the-efpia-view/statements-press-releases/efpia-statement-on-a-gdpr-code-of-conduct/)

Boehringer Ingelheim

universität wien
Department of Innovation and Digitalisation in Law

VirtualBrainCloud

WORKSHOP AGENDA

„Primary and secondary use of scientific research data“

22.6.2022, Boehringer-Ingelheim RCV Campus

(Altmannsdorfer Straße 55, 1120 Vienna)

09:00-17:00

09:00-09:15	Opening	Andrea Dillenz/ Michael Cepic
09:15-09:45	GDPR compliant Cloud Environments for Research on Human Digital Twins	Petra Ritter
09:45-10:15	How to balance data protection and scientific health research	Thomas Roth

Coffee break

10:45-11:15	GDPR and European Health Data Space: a first assessment	Nikolaus Forgó
11:15-12:00	Workshop introduction	All

Lunch break

13:00-14:30	Workshop groups	All

Coffee break

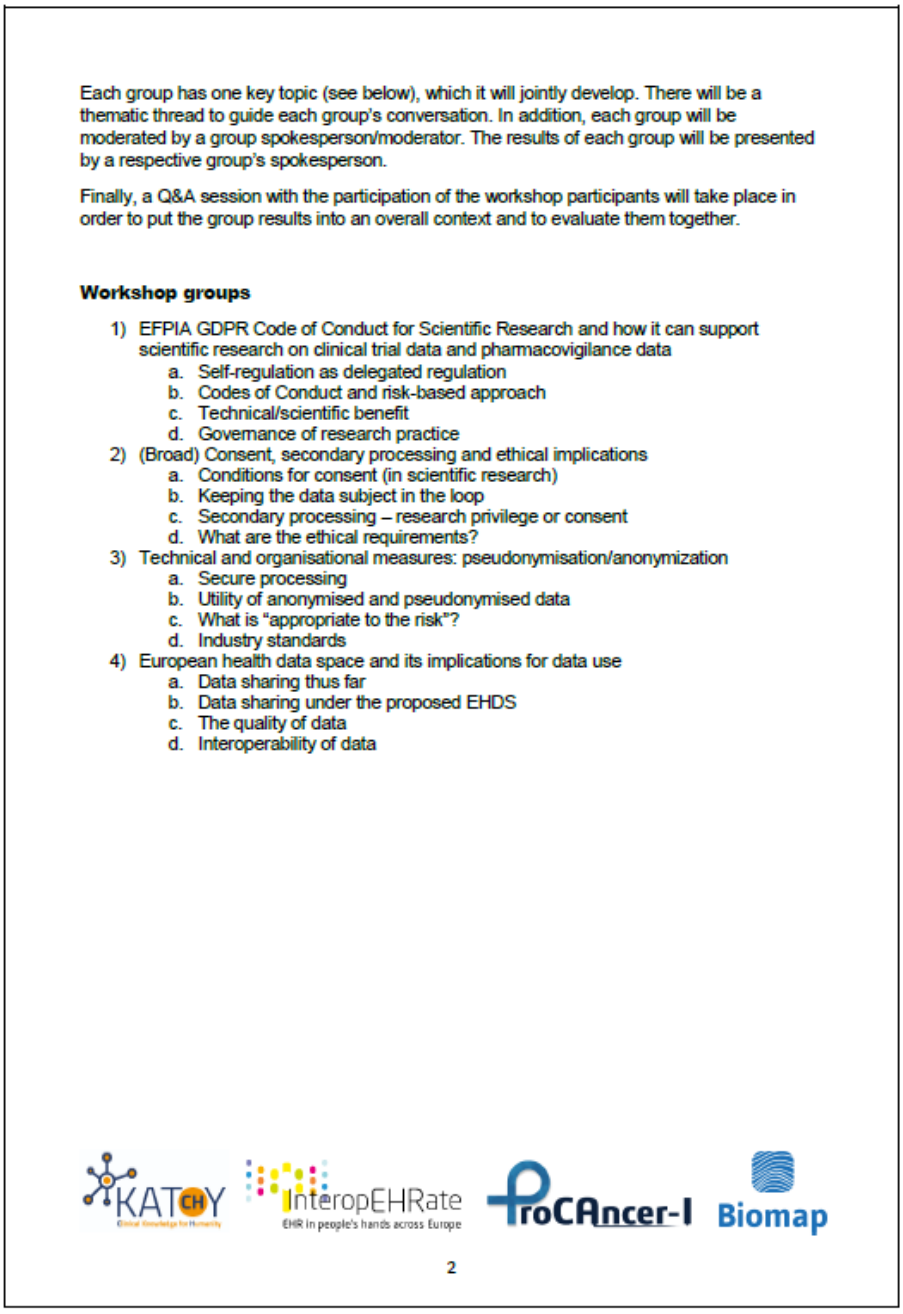
14:45-15:15	Workshop groups (conclusion)	All
15:15-15:45	Workshop groups (presentation)	All
15:45-16:45	Q&A	All
16:45-17:00	Closing	Andrea Dillenz/ Michael Cepic

Description

This workshop aims to bring together members of different disciplines in order to shed light on the theoretical and practical implementation of data protection requirements for data use in scientific research.

To this end, three presentations will first be given on the legal situation and practical application of data protection law in scientific research. Following this, four different topics will be dealt with interactively in small groups.

1



Each group has one key topic (see below), which it will jointly develop. There will be a thematic thread to guide each group's conversation. In addition, each group will be moderated by a group spokesperson/moderator. The results of each group will be presented by a respective group's spokesperson.

Finally, a Q&A session with the participation of the workshop participants will take place in order to put the group results into an overall context and to evaluate them together.

Workshop groups

1) EFPIA GDPR Code of Conduct for Scientific Research and how it can support scientific research on clinical trial data and pharmacovigilance data
   a. Self-regulation as delegated regulation
   b. Codes of Conduct and risk-based approach
   c. Technical/scientific benefit
   d. Governance of research practice
2) (Broad) Consent, secondary processing and ethical implications
   a. Conditions for consent (in scientific research)
   b. Keeping the data subject in the loop
   c. Secondary processing – research privilege or consent
   d. What are the ethical requirements?
3) Technical and organisational measures: pseudonymisation/anonymization
   a. Secure processing
   b. Utility of anonymised and pseudonymised data
   c. What is "appropriate to the risk"?
   d. Industry standards
4) European health data space and its implications for data use
   a. Data sharing thus far
   b. Data sharing under the proposed EHDS
   c. The quality of data
   d. Interoperability of data

KATCH InteropEHRate EHR in people's hands across Europe ProCAncer-I Biomap

2

(6) Organising an informational workshop on data sharing and privacy with national Alzheimer's Associations as part of an Alzheimer Europe Public Affairs meeting

(7) Organising a Stakeholder Conference on specific issues of neurodegenerative disease: the patients perspective, as part of the 2020 Alzheimer Europe conference (see D2.6)

Furthermore UNIVIE circulated among technical partners (EODYNE, CHARITE) a questionnaire which aimed to idetifiy not only challenges but also best practices and industry standards followed during the project. Questions and answers can be found in Annex I.

Whereas the replies were quite diverse, we provide below a summary of the mian findings: the main common denominators as follows:

Challenges regarding data protection legal framework encountered during the project:

- no challenge on personal data definition of brain images
- DPIA, lack of expertise of DPOs for their conduct
- Data sharing: no harmonized language of agremenets which poses a challenge to non-lawyers researchers and sharing data with counties with no EC adequacy decision
- Data minimization tampers with objective of future research projects
- FAIR

Best practices recommended by partners on the dat aproetction legal fraemwrok:

- (anonymization or pseudonymisation techniques) It was recommended to remove primary identifying information and treat the remaining data as sensitive health data that require thorough protection.

5. Conclusion

This Deliverable D2.5 provided a detailed overview of the legal and ethical frameworks applicable to the project while identifying best practices and lessons learned from the three last year of research and development of the TVB-Cloud.

With regard to the legal framework, data protection and cloud services related framework were identified. From a data protection perspective, not only general recommendations were provided to the partners but also specific questions related to the project were dealt with (are Brian images personal data?). Additionally, the cloud computer framework identified the main legal instruments that would apply to cloud services providers with particular reference to the security requirements as provided by law. Finally, yet importantly, two main prospective frameworks were identified which if in force will have an impact on the project regarding the obligations posed to the AI providers and to the data governance structure of the TVB-Cloud.

With regard to the ethical framework, we reflected on the ethical issues identified in D2.1 and discussed how these issues were raised over the duration of the project, identifying ways to address them in future development, and integrating the perspectives of caregivers through public involvement consultations. Five key ethical challenges with relevance to TVB-Cloud were identified: fairness & equity, informed consent, non-discrimination, confidentiality, and transparency. TVB-Cloud has developed a platform to accelerate brain research and support modelling of complex disease processes using brain simulations

built on multimodal data. As an open platform that is an enabler for research, TVB-Cloud can enhance fairness and equity for scientists in resource-poor settings. Measures to consider going forwards should further improve access, awareness and understanding for researchers and clinicians, and for key beneficiaries of TVB-Cloud; patients, research participants and caregivers. In particular, these measures should consider structural inequalities and diverse sources of bias (from algorithmic bias, to bias in access to AI to reduce the risk of widening existing healthcare disparities, and ensure TVB-Cloud can benefit all groups within society.

Lastly, lessons learned were identified and main challenges on the compliance of GDPR were identified.

Annex I

Deliverable 2.5 - best practices in VBC – questions for the partners - Charité

1. Implementation of identified data protection requirements in the health research such as the VBC:

    a. Defining personal data – what were the main challenges in transposing the GDPR definition of personal data in practice and how were they overcome? Is it possible, considering available technology, to identify an individual based on his/her brain image? What is the conclusion from a technical perspective?

Transposing the GDPR definition of personal data in practice was not considered a challenge.

To identify an individual based on his/her brain image is possible.

Radiological brain images are considered personal health data.

    b. What techniques have been used to anonymise or pseudonymise personal data of research participants? Are there any recommendations which techniques are the most recommended?

For personalized brain simulations data are not anonymized b/c they would lose the information subject to the actual research purpose.

Radiological brain images of humans cannot be considered pseudonymized after removing primary identifying information such as name, birthdate, address b/c identifying features from the brain can be still used for re-identification.

We recommend to remove primary identifying information and treat the remaining data as sensitive health data that require thorough protection.

    c. Data protection impact assessment – were there any challenges in that regard and how were they overcome?

In my experience the biggest challenge is the lack of domain specific expertise / knowledge / competence of the relevant DPOs.

d. Data sharing contracts – what technical and governance aspects must be considered in a health research project such as TVB-C, when defining the roles of controllers and processors?

The challenge is to set up the sharing agreements with proper legal language given that scientists are no lawyers and standard templates/language for such agreements are still lacking. The actual definition of roles (controller, joint controllers, processor) is straightforward.

e. Storage limitation and technical solutions enabling appropriate storage and data minimization – are there any suggestions from a technical perspective that in your view projects similar to TVB-C should consider?

Data minimization stands in a general conflict with the desire to keep health data for future research projects and hence store comprehensive data sets rather than limiting data to what is needed for a single research question and discarding all data not directly needed for that.

f. Best practices in data FAIR-ification – what are the main lessons learnt from the technical and scientific perspective? E.g. data interoperability, data access.

Many aspects need to be considered, including the processes for provenance tracking, annotation, data ingestion in platforms etc. Making data FAIR is a very comprehensive task.

g. What technical and organizational measures are implemented in TVB-C to protect personal data?

Access control, encryption, sandboxing, Iso 27001 standards, information security certification of the IT provider, shared liability model with the corresponding agreements between stakeholders

h. Automated decision-making – what solutions have been implemented in TVB-C? Is such automated decision-making, that involves personal data concerning health, based on the consent of the data subject?

All processing of personal data is based on consent of data subjects.

Presently no decision making for clinical practice is made by TVB_-Cloud.

i. What challenges where faced when sharing personal data for research purposes?

Sharing/using data from outside EU b/c of lack of GDPR compatible consent ensuring subject rights as of GDPR. Use of historic data acquired before GDPR was in force. Sharing data to other countries with no GDPR adequacy agreement.

2. Owner/co-ownership of TVB-Cloud environment – what technical aspects have had to be taken into account when assessing the ownership of developed TVB-Cloud environment?

IP ownership

3. Cloud computing – best practices in the development of the TVB-Cloud:

    a. What data security/architectural aspects were important to be considered in the establishment of such an infrastructure? Please consider for e.g.:

        i. Technical and organizational security measures:

            1. Pseudonymisation/anonymization – lower importance
            2. Data minimization – still a difficult concept / see above
            3. Solutions enabling continuous confidentiality, integrity and availability and resilience of processing systems and services – high priority and solutions in place
            4. Procedures established in case of physical or technical incident – high priority and solutions in place
            5. Regular testing, assessing and evaluating effectiveness of technical and organizational measures - high priority and solutions in place

        ii. Agreements with cloud service providers (if applicable) and arrangements in the case of transfer of personal data to the third countries (outside EEA)

No cloud service providers are presently being used. While EU providers in principle can be considered, providers form outside EU are not an option due to their national law contradicting privacy regulations in EU.

        iii. Risk management

Subject to DPIA

iv. Cloud architecture and privacy by design

Has been developed as part of VRE

v. Adherence to any security standards (e.g. ISO 27001)

ISO 27001 and BSI

b. Integration with the VRE – what aspects have been important for the achievement of compliance with data protection regulations? (Based on use case of D 3.5 integration to the VRE).

VRE provides the audited GDPR compliant platform for TVB-Cloud.

c. What industry standards have been considered in the VBC development phase and in the integration with the VRE?

ISO27001 and BSI

d. Access Control Management – are the access rights limited? What were the main challenges in implementation of appropriate access management strategy?

Integration with Charite active directory. External users undergo identification and conclude a contractual agreement with Charite.

e. Authentication and authorization – what one should consider when structuring these elements in the cloud solution such as TVB-Cloud?

TVB-Cloud uses AAI infrastructure of Charite. In future several AAI will be federated, e.g. the one of EBRAINS.

f. What network connections are considered secure for such infrastructure?

The VRE platform is hosted behind the Charité firewall. All user interactions with the data and platform content must take place via the research portal or VRE-managed command line or workbench utilities. The platform does not provide users with direct access or connection to the underlying resources such as storage, databases, services, or VMs.

All services, including the API Gateway, are inaccessible from outside of the VRE deployment. The Ingress Controller defines rules for external connectivity, e.g., from the Charité out-facing proxy, to the VRE deployment, including URL rewriting, upstream services, etc.

A single-entry point situated between the Ingress Controller and VRE platform services. The API Gateway is responsible for request routing, composition, rate limitation, and monitoring. It handles requests by routing them to the appropriate back-end services. If there are failures in the back-end services, the API Gateway can return cached or default data. In the VRE deployment, the API Gateway is also connected to the internal Identity Provider (IdP) to protect the back-end APIs from unauthorized access.

The VRE service components - the front-end, API Gateway that connects all back-end services, and workbench tools - are registered into Keycloak as individual clients using OpenID Connect (OIDC) as the authentication protocol. The typical authentication workflow in VRE is described as the follows: VRE portal asks the Keycloak to authenticate a user. After a successful login, VRE portal receives an access token that contains username, email, other profile information, and access details such as role mapping. The access token is digitally signed by Keycloak and can be used by other registered clients to invoke other services on behalf of the user. The service that receives the request then extracts the access token, verifies the signature of the token, and decides based on access information within the token whether to process the request.

g. Collaboration with external cloud service provider – what are the main lessons learnt from this collaboration?

None

4. AI – best practices in the development of the TVB-Cloud:

   a. What were the main challenges in the creation of the AI solutions and their implementation?

None. AI solutions can run in TVB-Cloud as any other complex analyses or simulations.

Deliverable 2.5 - best practices in VBC – questions for the partners - Eodyne

1. Implementation of identified data protection requirements in the health research such as the VBC:

   a. Defining personal data – what were the main challenges in transposing the GDPR definition of personal data in practice and how were they overcome? Is it

possible, considering available technology, to identify an individual based on his/her brain image? What is the conclusion from a technical perspective?

In the case of Eodyne does not deal with brain images. The RGS system collects Kinematics movements from patients.

The main challenges in transposing the GDPR definition of personal data were on maintaining full anonymization of all the data and profiles in the system and databases.

b. What techniques have been used to anonymise or pseudonymise personal data of research participants? Are there any recommendations which techniques are the most recommended?

To capture data from research participants all the registration and data collection in the RGS system is anonymized.

c. Data protection impact assessment – were there any challenges in that regard and how were they overcome?

d. Data sharing contracts – what technical and governance aspects must be considered in a health research project such as TVB-C, when defining the roles of controllers and processors?

e. Storage limitation and technical solutions enabling appropriate storage and data minimization – are there any suggestions from a technical perspective that in your view projects similar to TVB-C should consider?

We considered the storage limitation when we designed the format of the log files, for example, by reducing data redundance and precision.

f. Best practices in data FAIR-ification – what are the main lessons learnt from the technical and scientific perspective? E.g. data interoperability, data access.

g. What technical and organizational measures are implemented in TVB-C to protect personal data?

Eodyne has a personal data expert lawyer that advises the company on the processes that has to be implemented to follow the regulations.

h. Automated decision-making – what solutions have been implemented in TVB-C? Is such automated decision-making, that involves personal data concerning health, based on the consent of the data subject?

i. What challenges where faced when sharing personal data for research purposes?

The RGS system already implements meausres to keep data anonymized and there was no additional challenge to use it for research purposes.

2. Owner/co-ownership of TVB-Cloud environment – what technical aspects have had to be taken into account when assessing the ownership of developed TVB-Cloud environment?

3. Cloud computing – best practices in the development of the TVB-Cloud:

a. What data security/architectural aspects were important to be considered in the establishment of such an infrastructure? Please consider for e.g.:

i. Technical and organizational security measures:

1. Pseudonymisation/anonymization
2. Data minimization
3. Solutions enabling continuous confidentiality, integrity and availability and resilience of processing systems and services
4. Procedures established in case of physical or technical incident
5. Regular testing, assessing and evaluating effectiveness of technical and organizational measures

ii. Agreements with cloud service providers (if applicable) and arrangements in the case of transfer of personal data to the third countries (outside EEA)

iii. Risk management

iv. Cloud architecture and privacy by design

v. Adherence to any security standards (e.g. ISO 27001)

b. Integration with the VRE – what aspects have been important for the achievement of compliance with data protection regulations? (Based on use case of D 3.5 integration to the VRE).

c. What industry standards have been considered in the VBC development phase and in the integration with the VRE?

d. Access Control Management – are the access rights limited? What were the main challenges in implementation of appropriate access management strategy?

e. Authentication and authorization – what one should consider when structuring these elements in the cloud solution such as TVB-Cloud?

f. What network connections are considered secure for such infrastructure?

g. Collaboration with external cloud service provider – what are the main lessons learnt from this collaboration?

4. AI – best practices in the development of the TVB-Cloud:

a. What were the main challenges in the creation of the AI solutions and their implementation?

The challenge of the development of AI solutions is on the validation. In the case of RGS, we have created an adaptive difficulty component based on the performance of the patient during rehabilitation activities.

Annex II

Annex to D3.4 on the Use Case Scenario for personal data sharing ('Interim source-space multiresolution MEG time series in BIDS share. Initial MEG and SEEG brain dynamic measures at the disposal of other WPs') Contribution from UNIVIE to D3.4

1. Objective

Personal data (hereinafter, 'personal data' and 'data' will be used interchangeably) sharing is the process of making data available to others. For health research purposes, this frequently involves personal data, where a person is identified or identifiable. As a result, the data sharing process must be in line with data protection laws, including the EU General Data protection Regulation ('GDPR')[176].

Next to legal requirements, it is mostly agreed upon, that the burdens and benefits of data sharing should be fairly allocated, trying to achieve what is called a 'responsible data sharing'. Benefits for the individual and the society should be maximised while harms minimized. Complying with legal requirements will indicate overall conformity with society's principles.[177]

2. Purpose

Using the use case described in D3.4, the purpose of this use case scenario is to demonstrate steps that are necessary to establish GDPR-compliant personal data sharing for research purposes of the TVB-Cloud among TVB-Cloud partners during the duration of the Project. It is important to highlight that this use case should serve as a basis for future data sharing within the TVB-Cloud Project.

It is not only required, but might prove profitable for the future, if the data protection laws are minded from the beginning of the TVB-Cloud Project,[178] a fact that is being implemented in TVB-Cloud Project. In order to support this argument, Art. 25 GDPR requires for a data protection-by-design approach "*at the time of the determination of the means for processing and at the time of the processing itself*". As a result, controllers[179] need to implement the required measures at the final stage of the planning of their system.[180] This also include the data sharing activities planned for a project.

3. Method

[176] 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).

[177] *Kalkman* et al., Responsible data sharing in international health research: a systematic review of principles and norms, BMC Medical Ethics 2019, 20:21.

[178] *Gruschka* et al., Privacy issues and data protection in big data: A case study analysis under GDPR.

[179] 'Controller' is defined in Article 4(7) GDPR as 'the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data; where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law.

[180] *Mantz*, Art 25, in *Sydow* (Hrsg), Europäische Datenschutzgrundverordnung (2017) 604 (615).

Being originally developed to describe functional requirements of software systems from a user perspective, the use case approach can be helpful in order to understand any given process of a (software) system[181]. Even more, the use case driven approach is something quite common in data security research. To identify and analyse relevant legal requirements it is necessary to assess the specific data sharing process[182]. For this reason a step-by-step analysis and description of the TVB-Cloud D3.4 task "Interim source-space multiresolution MEG time series in BIDS share. Initial MEG and SEEG brain dynamic measures at the disposal of other WPs data sharing processes fits best, to map out legal issues that arise" was performed.

UNIVIE's method to construe the use case was to distribute guiding questions to the partners involved (UH, UCM and UNIGE) and gather information regarding the intended personal data sharing for the purpose of the task developed in the framework of D3.4. In the form of a Q&A the questions below were asked to the partners:

1. Which dataset do you intend to share?
2. What type of data? Are these data identifiable? Is it planned to share all data contained in the dataset?
3. Which partner has in their possession the dataset which is intended to be shared?
4. To which partner(s) is it planned for the data to be shared?
5. Has the partner who possess the data a legal basis to share such data? Do you have the legal authority to share it?
6. Has the partner who possess the data require ethics approval?
7. Are there any legal conditions/restrictions for the use/sharing of the dataset? Are there any restrictions to process it? If yes, which ones?
8. How are you planning to share it (transmission)? Which security measures are required to follow/will be implemented?
9. How is the dataset going to be stored? At the recipient's? In a central repository?
10. How is the dataset going to be stored? At the recipient's?
11. Which security measures are you going to use for the protection of data (transm./storage)?
12. What data management tool are you using for recording the actions you do with datasets?
13. Did the partner having the dataset submit the statement of DPO?
14. What role do you have in regard to the cohort? Are you making a decision on the processing of data? Are you deciding on the purpose of how to process such data? Are you deciding on the means of the data processing? Are you doing this jointly with other partner(s)? Are you

[181] *Egan* et al., Compilation of food composition data sets: an analysis of user needs through the Use Case approach, European Journal of Clinical Nutrition 2011, 757 (758).

[182] Cf. *Mai* et al., Modeling security and privacy requirements: a Use Case-driven approach, Information and Software technology, 165 (166).

following instructions of another partner to process data? Are you making decisions on the processing of data jointly with other partners? Are you interested in the result of the processing? How long are you going to process the data?

Answers to the questions mentioned above and thus the use case based information can be found in point 6 below.

Partners will also need to assess whether they are controllers or processors. Therefore, please tick the corresponding boxes below. This **checklist**[183] sets out indicators as to whether you are a controller, a processor or a joint controller on the personal data you are planning to receive from another partner of the TVB-Cloud. The more boxes you tick, the more likely you are to fall within the relevant category.

Are we a controller?
☐ We decided to collect or process the personal data.
☐ We decided what the purpose or outcome of the processing was to be.
☐ We decided what personal data should be collected.
☐ We decided which individuals to collect personal data about.
☐ We obtain a commercial gain or other benefit from the processing, except for any payment for services from another controller.
☐ We are processing the personal data as a result of a contract between us and the data subject.
☐ The data subjects are our employees.
☐ We make decisions about the individuals concerned as part of or as a result of the processing.
☐ We exercise professional judgement in the processing of the personal data.
☐ We have a direct relationship with the data subjects.
☐ We have complete autonomy as to how the personal data is processed.
☐ We have appointed the processors to process the personal data on our behalf.

Are we a joint controller?
☐ We have a common objective with others regarding the processing.
☐ We are processing the personal data for the same purpose as another controller.
☐ We are using the same set of personal data (eg one database) for this processing as another controller.

[183] This checklist was taken from ICO https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/key-definitions/controllers-and-processors/, accessed 29. November 2022.

☐ We have designed this process with another controller.

☐ We have common information management rules with another controller.

Are we a processor?

☐ We are following instructions from someone else regarding the processing of personal data.

☐ We were given the personal data by a customer or similar third party, or told what data to collect.

☐ We do not decide to collect personal data from individuals.

☐ We do not decide what personal data should be collected from individuals.

☐ We do not decide the lawful basis for the use of that data.

☐ We do not decide what purpose or purposes the data will be used for.

☐ We do not decide whether to disclose the data, or to whom.

☐ We do not decide how long to retain the data.

☐ We may make some decisions on how data is processed, but implement these decisions under a contract with someone else.

☐ We are not interested in the end result of the processing.

4. Results

Reviewing the partners' feedback allows us to identify the data type that is shared, including possession, and sharing target. Furthermore, we can assess the legal basis, processing and sharing restrictions and possible conditions. As to storage and security, we can establish storage location, security measures and their implementations as well as data management tools. Finally, we can describe the processing activities, including purpose, joint or individual processing and results.

In this regard, there are two datasets being shared: Magnetoencephalography (MEG) data and Stereoelectroencephalography (SEEG) data. The cohort information should result in this from per cohort:

4.1 Cohort Summary: UCM Magnetoencephalography (MEG) data

The information provided by the partners is summarized below in the following categories:

- Data type/identifiability: The data contained therein is either processed pseudonymized data with numeric codes (unique for the TVB-Cloud project) or statistical derivatives of pre-collected raw data. UCM is in possession of the MEG dataset.

- Sharing activity: MEG (processed data and derivatives) will be shared among University of Helsinki, Charité, FZJ and AMU and other WP3 and WP8 partners **who are not involved in this deliverable but might need data at later stages**.
- Legal basis for sharing: With regard to MEG data a legal basis was confirmed by UCM's DPO, which is in line with the Spanish data protection law (Ley Orgánica de Protección de Datos, LOPD).
- Legal basis for processing data: The processing of previously collected data for further scientific use has been accepted by participants of the studies conducted by the partners by providing signed informed consent and agreed upon on the corresponding previous ethical committees.
- Legal restrictions: There is no sharing of raw unprocessed data. As to MEG for some of the recordings, the origin of the dataset is the UCM. Data were pooled from the databases of two projects of the UCM supported by the Spanish Ministry of Science and Innovation (PSI2009-14415-C03-01) and the Spanish Ministry of Economy and Competitiviness (PSI2012-38375-C03-01). Our agreement with third party does not affect the sharing of this data...
    - Storage: MEG will be available through an UCM server accessible at the URL <https://vbc.ucm.es/login.php>, where data is stored.
- Security measures: Security levels regarding the storage in the UCM are guaranteed in compliance with Spanish and European law, as established after an initial external audit and audited every 6 months by the DPO. For the transmission of data, the UCM server described above will be used. To keep track of the downloads, the UCM server integrates the above-described LOG system which allows for automatic tracking of the filename, solicitant name, and timestamp for any file that has been requested for download (Article 32 GDPR)
- Data management: For both datasets, a record of data recipient/data description/exchange date will be kept in a spreadsheet to maintain control over the manual file exchange. Additionally, for MEG data, the LOG system described above will keep track of the data flow through the UCM server. (Article 30 GDPR)
- Controllership: Decisions regarding MEG data are so far made by UCM, which will indicate to the recipients TVB-Cloud partners for which purpose an how the processing will take place. .

From the information above, UNIVIE was able to identify the elements relevant for the establishment of a DSA, if needed:

a) Original Data Controller: UCM
b) Recipient TVB-Cloud Partner: UCM
c) Legal Basis for processing data(e.g. licensing agreement between a &b): N/A
d) Desired recipient TVB-Cloud Partners: University of Helsinki, Charité, FZJ and AMU.
e) Legal Basis for sharing data with consortium: DPO letter of confirmation to share data.
f) Need of DSA: YES, among UCM and University of Helsinki, Charité, FZJ and AMU. Template agreement suggested for this scenario is in point 7.

1.2 Cohort Summary: SEEG data

The information provided by the partners is summarized below in the following categories:

- Data type/identifiability: We will share Functional connectivity measures (PLV, PAC, CFC), dynamical phase lags, and criticality indices (DFA and bistability). All these statistical values are going to be estimated between all possible brain signals pairs extracted from SEEG invasive electrodes. We will also share meta-data information such as contact positions within single-subject brain space, most proximal anatomical areas to each contact. These are saved as binary matrices in Matlab- and python-compliant formats. Note that SEEG data cannot be shared because our current agreement do not cover that (we are not allowed to share raw data outside the network of collaboration that involve UNIGE-UH-Niguarda). This dataset is a set of retrospective SEEG data, collected at Niguarda Hospital in accordance to a scientific collaboration agreement. Raw data (i.e. sensor time-series) will not be shared. As defined in the Data Sharing Agreement, we will share statistical derivatives (see above) for each of the subjects involved in the study. In our view, our results represent an invaluable resource for optimal optimization of model parameters. Given the unprecedented observations of long-range phase synchronization profiles, the submillimeter accuracy of spatial mapping, the phase-dynamics profiles, VBC modeling groups could take into account our results and by opportune model parameter optimization try to accomodate our physiological observations in non-demented controls. These will serve to establish with meso-scale recordings the properties of the “healthy” aging brain.
- Sharing activity: SEEG datasets will be shared among University of Helsinki, Charité and other WP3 and WP8 partners determined by FZJ (Lead WP3) and AMU (Lead WP8). Additionally, the SEEG datasets will also be shared with a member of WP4.
- Legal basis for sharing: There is an agreement being negotiated between Niguarda Hospital and UniGe and UH. The agreement which we currently have allows us to share the results of our analyses but not the raw data.
- Legal basis for processing data: The processing of previously collected data for further scientific use has been accepted by participants of the studies conducted by the partners by providing signed informed consent and agreed upon on the corresponding previous ethical committees.
- Legal restrictions: According to the terms of the draft agreement with Niguarda Hospital, derived data of SEEG could be shared without restrictions.
- Storage: For Interim version of the UNIGE-UH datasetSEEG, we can envisage two possible ways according to current UNIGE infrastructures. Either setting up a local server with restricted access policies to which the receiving partner will be given access, or, using some cloud service that is GDPR compliant (e.g. Dropbox). However, it would be best to use VBC infrastructures or even EBrains.

- Security measures: Regarding the SEEG data the raw data are saved in secure server in UNIGE maintained with high-security levels and complying with current national and European regulations. (Article 32 GDPR)
- Data management: For both datasets, a record of data recipient/data description/exchange date will be kept in a spreadsheet to maintain control over the manual file exchange. Additionally, for MEG data, the LOG system described above will keep track of the data flow through the UCM server. (Article 30 GDPR)
- Controllership: Regarding SEEG data decisions on data are jointly made between the University of Helsinki and Niguarda Hospital.. (Article 26 GDPR).

From the information above, UNIVIE was able to identify the elements relevant for the establishment of a DSA, if needed:

a) Original Data Controller: Niguarda Hospital
b) Recipient TVB-Cloud Partner: UNIGE and UH
c) Legal Basis for processing data(e.g. licensing agreement between a &b): Draft data sharing agreement between Niguarda hospital and UNIGE
d) Desired recipient TVB-Cloud Partners: University of Helsinki, Charité, FZJ and AMU
e) Legal Basis for sharing data with consortium: DPO confirmation statement.
f) Need of DSA: NO. The type of data planned to be shared between UNIGE and UH with other TVB-Cloud partners (Functional connectivity measures (PLV, PAC, CFC), dynamical phase lags, and criticality indices (DFA and bistability) is not considered to be personal data, according to the definition provided by the GDPR, thus the GDPR is not applicable, and therefore from a data protection perspective, no agreement needs to be to put in place.

5. Summary / Implications for Partners

The steps to be taken by partners in the TVB-Cloud project before sharing personal data are the following:

1. Identify the internal data flow and map out the data sharing origin and target. Reply to the Q&A drafted by UNIVIE prior to any data sharing activity. The Q&A serves as a guidance to map the internal data flow.
2. Identify data(sets) and data types (personal, non-personal, sensitive, health, genetic, biometric), map categories of data.
3. Establish the legal basis for data sharing and data sharing restrictions and/or conditions of the data originator/data custodian/ partner in possession of personal data, if any.
4. Assess the processing activity(-ies), including purpose, length, possible results of processing as well as the partners involved and their roles.
5. Describe security measures for data sharing.
6. Discuss and establish storage and data sharing duration.
7. Record in written point 1-6.

8. Identify the need for a Data sharing Agreement. A template of a Controller-Processor Data Sharing Agreement can be found in point 7. It is suggested that UNIVIE is consulted in regard to the type of Data Sharing Agreement to be put in place.

6. Q&A Reply from Partners

The questionnaire copied below was circulated on 17 January 2020 by UNIVIE to WP3 partners and the reply from WP3 was received by UNIVIE on 11 February 2020.

TVB-CLOUD – DATA SHARING STEPS – WHAT DO WE NEED IN ORDER TO SHARE? (NON-EXHAUSTIVE LIST OF QUESTIONS) – Use case – D3.4

The questions for your completion are below. It would be really helpful for us if you could expand your answers. The more information we have, the better is for us to assess data protection implications. Thank you very much. Replies from partners UH, UCM & UNIGE:

Q: Which dataset do you intend to share?
There are two kinds of datasets. One from UCM and one from UNIGE-UH. The UCM dataset is oriented to Magnetoencephalography (MEG) data and UNIGE-UH dataset is oriented to Stereo-Electroencephalography (SEEG) data. In D3.4, we intend to share interim versions of the two datasets as described below. **Interim version of the UCM dataset** Processed MEG data for 20 participants (healthy controls (n = 5), participants with Subjective Cognitive Decline (n = 5), participants with Mild Cognitive Impairment (n = 5), and participants with Alzheimer's Disease (n = 5)). MEG data includes source-space time-series and derivatives from the MEG pre-processing and source reconstruction pipelines (e.g. artifact-cleaned and epoched MEG data, leadfield, headmodel, sourcemodel, beamformer filters, etc). Additionally, we provide T1-weighted MRI data, as well as relevant genetic data (APOE alleles) for each participant. Functional connectivity measures (both static and dynamic) are also provided. **Interim version of the UNIGE-UH dataset** We will share Functional connectivity measures (PLV, PAC, CFC), dynamical phase lags, and criticality indices (DFA and bistability). All these statistical values are going to be estimated between all possible brain signals pairs extracted from SEEG invasive electrodes. We will also share meta-data information such as contact positions within single-subject brain space, most proximal anatomical areas to each contact. These are saved as binary matrices in Matlab- and python-compliant formats. Note that SEEG data cannot be shared because our current agreement do not cover that (we are not allowed to share raw data outside the network of collaboration that involve UNIGE-UH-Niguarda). Extending that agreement will take time, politics sufferings and ultimately it also requires reaching back **all** patients to sign a different agreement. Final and most important part those data (ie raw SEEG series) were never mentioned in the data agreement we signed.
Q: What type of data? Are these data identifiable? Is it planned to share all data contained in the dataset?
Interim version of the UCM dataset **Pseudonymized data**: Upon enrolling, participants were assigned a **numeric code** provided at the medical center where recruitment was carried out. No personal data (name, date of birth, personal ID, personal address,etc) were collected during the experimental procedure at the UCM; only general demographics. The MEG data were labelled using the assigned numeric code. This original numeric code was converted to a new ID (different to the one assigned at the medical center and unique for the TVB-Cloud project) according to BIDS practice (i.e. sub-XXX). The only link between the original

(not provided to partners) numeric code and the personal data (the name and signature of the participants) would be found in the signed informed consents that are safely stored in compliance with National and European Laws. Additionally, T1-weighted MRI data have been defaced to remove identifiable features from the images. MEG raw data is not to be shared, only data that has already been processed.

Interim version of the UNIGE-UH dataset

This dataset is a set of retrospective SEEG data, collected at Niguarda Hospital in accordance to a scientific collaboration agreement. Raw data (i.e. sensor time-series) will not be shared. As defined in the Data Sharing Agreement, we will share statistical derivatives (see above) for each of the subjects involved in the study. In our view, our results represent an invaluable resource for optimal optimization of model parameters. Given the unprecedented observations of long-range phase synchronization profiles, the submillimeter accuracy of spatial mapping, the phase-dynamics profiles, VBC modeling groups could take into account our results and by opportune model parameter optimization try to accomodate our physiological observations in non-demented controls. These will serve to establish with meso-scale recordings the properties of the "healthy" aging brain.

Q: Which partner has in their possession the dataset which is intended to be shared?

The **interim version of the UCM dataset** is in the possession of the UCM.
Interim version of the UNIGE-UH dataset is in possession of the UNIGE and UH.

Q: To which partner(s) is it planned for the data to be shared?

Both datasets are planned to be shared with University of Helsinki, Charité, and relevant partners from WP3 and WP8.

Q: Has the partner who possess the data a legal basis to share such data? Do you have the legal authority to share it?

Interim version of the UCM dataset

Yes. The Data Protection Officer (DPO) of UCM has given us the approval to share the data following the requirements specified by the Spanish Data Protection Law (Ley Orgánica de Protección de Datos, LOPD). The signature of confidentiality agreements with each partner making use of the shared dataset will be required in order to comply with DPO guidelines (we have the template available).Yes, we have the legal basis to share the data. The data protection office of UCM has given us the approval to share the data, following the requirements specified by the Spanish data protection law (LOPD). The signature of a confidentiality agreement by each partner making use of the shared data is required in order to comply with the DPO guidelines.

Interim version of the UNIGE-UH dataset

To be discussed with DPO

Q: Has the partner who possess the data require ethics approval?

For both datasets:

No, the processing of previously collected data for further scientific use has been accepted by participants by providing signed informed consent and agreed upon on the corresponding previous ethical committees.

Q: Are there any legal conditions/restrictions for the use/sharing of the dataset? Are there any restrictions to process it? If yes, which ones?

Interim version of the UCM dataset

Raw unprocessed data is not to be shared. For some of the recordings, we have data sharing or licensing agreements with third parties to the TVB-Cloud project (e.g. the BioFIND project (JPND research)) that do not affect the sharing of processed data within the consortium.

Interim version of the UNIGE-UH dataset

Raw unprocessed data is not to be shared. Derived data could be and would be shared.

Q: How are you planning to share it (transmission)? Which security measures are required to follow/will be implemented?

Interim version of the UCM dataset

Data are available through a **UCM server** accessible at the Uniform Resource Locator (URL) https://vbc.ucm.es/login.php. This data will be available for research purposes within the TVB-Cloud project. Upon request, a login will be created for the solicitant who will then be able to access the data and metadata under IP approval. In order to request access, the solicitant must specify **Name, Surname, Institution and TVB-Cloud Work Package.** Access rights to third parties outside TVB-Cloud will be denied. **A record of requests** (*data recipient/data description/exchange date)* will be automatically generated on an electronic document to maintain control over the data flow from the server (LOG system). If request came from unanticipated partners (i.e. those not expected to process data in any way) further information regarding purpose and objectives will be demanded. To further secure data exchange, **bilateral agreements** between UCM and partners with whom data is being shared should be arranged as soon as possible. For **Interim version of the UNIGE-UH dataset,** We can envisage two possible ways according to current UNIGE infrastructures. Either setting up a local server with restricted access policies to which the receiving partner will be given access, or, using some cloud service that is GDPR compliant (e.g. Dropbox). However, it would be best to use VBC infrastructures or even EBrains.
Q: How is the dataset going to be stored? At the recipient's? In a central repository?
The Interim version of the UCM datasetdata is available through an **UCM server** accessible at the URL https://vbc.ucm.es/login.php — where data is stored. **Interim version of the UNIGE-UH dataset** Depending on the chosen scenario in answer above, we can either set up a local repository for data or share a central repository.
Q: How is the dataset going to be stored? At the recipient's?
The **Interim version of the UCM dataset** is available through an **UCM server** accessible at the URL **https://vbc.ucm.es/login.php** where data is stored. The **Interim version of the UNIGE-UH** raw pseudonymized as well as processed data are stored in a Network Archiving System (NAS) within the UniGe (DIBRIS) LAN and is accessible only from within the institutional network (or remotely accessed via VPN).
Q: Which security measures are you going to use for the protection of data (transm./storage)?
Interim version of the UCM dataset Security levels regarding the storage in the UCM are guaranteed in compliance with Spanish and European law, as established after an initial external audit and audited every 6 months by the DPO. For the transmission of data, the UCM server described above will be used. To keep track of the downloads, the UCM server integrates the above-described LOG system which allows for automatic tracking of the filename, solicitant name, and timestamp for any file that has been requested for download. **Interim version of the UNIGE-UH dataset** The raw data are saved in secure server in UNIGE maintained with high-security levels and complying with current national and European regulations.
Q: What data management tool are you using for recording the actions you do with datasets?
For both datasets, a record of *data recipient/data description/exchange date* will be kept in a spreadsheet to maintain control over the manual file exchange. Additionally, for the UCM data, the LOG system described above will keep track of the data flow through the UCM server.
Q: Did the partner having the dataset submit the statement of DPO?
Yes, the statement of DPO has been submitted for both datasets.
Q: Are you making a decision on the process of data?
Interim version of the UCM dataset Decisions are so far made jointly between UCM and the University of Helsinki. **Interim version of the UNIGE-UH dataset** Decisions regarding SEEG data are so far made jointly between the University of Helsinki and Niguarda Hospital

Q: Are you deciding on the purpose of how to process such data? Are you doing this jointly with other partner(s)?
Decisions regarding MEG data are so far made jointly between UCM and the University of Helsinki, and regarding SEEG data, between the University of Helsinki and Niguarda Hospital.
Q: Are you deciding on the elements of the processing? Are you doing this jointly with other partner(s)?
Decisions regarding MEG data are so far made jointly between UCM and the University of Helsinki are jointly deciding on the elements of processing so far for MEG data, and joint decision are made between the University of Helsinki and Niguarda Hospital for SEEG data.
Q: Are you following instructions of another partner to process data?
No, decisions regarding MEG data are so far made jointly between UCM and the University of Helsinki. For SEEG data, the University of Helsinki and Niguarda Hospital are jointly taking decisions.
Q: Are you making decision on the process of data jointly with other partners?
Yes, decisions regarding MEG data are so far made jointly between UCM and the University of Helsinki, and University of Helsinki and Niguarda Hospital jointly making decision for SEEG data.
Q: Are you interested in the end result of the processing?
Yes. We should or would like to be involved in the studies using these datasets (separately or combined).

Thank you.