945358-BIGPICTURE

Central Repository for Digital Pathology

WP5 - Regulatory framework for digital slides and AI-based methods

D5.01 Overview of the legal frameworks of all countries involved in BIGPICTURE

Lead contributor	35 - TIME.LEX
	Magdalena.Kogut@timelex.eu; Jos.Dumortier@timelex.eu
Other contributors	2 - LINKÖPING UNIVERSITY
	18 - BIOBANKS AND BIOMOLECULAR RESOURCES RESEARCH INFRASTRUCTURE CONSORTIUM affiliated entities: KU Leuven, UNIMIB and TURUN
	20 - OSTERGOTLANDS LANS LANDSTING
	22 - EUROPEAN INSTITUTE FOR INNOVATION THROUGH HEALTH DATA
	24 - PHILIPPS UNIVERSITAET MARBURG
	31 - STICHTING LYGATURE

Document History

Version	Date	Description
V0.0.1	1/9/2021	Outline of the report
V0.1	11/10/2021	First draft
V0.2	14/10/2021	Initial comments from internal review
V0.3	14/10/2021	Intermediate draft for review of Project Management
V0.4	18/10/2021	Draft for review of the Management Board
V0.5	28/10/2021	Draft with changes recommended by the Management Board
V1.0.	29/10/2021	Final Version

Table of Contents

Abstract

The objective of this report is to identify legal rules applicable to the initial contributors (Data Submitters; for this and further definitions and abbreviations, please refer to Glossary and Abbreviations section below) of Whole Slide Images ("**WSI**") and accompanying Metadata in the BIGPICTURE project in the scope that those rules may impact their provision and use within the project. The initial Data Submitters of the WSIs are located in the following 8 jurisdictions: Austria, Belgium, Finland, Germany, Italy, the Netherlands, Sweden and UK and the report outlines the main areas of legal interest on EU and national level. Those areas include: protection of personal data, laws applicable to non-clinical WSIs (animal WSIs), ethical requirements, rules relating to biobanks, intellectual property ("**IP**") protection, open data and contractual limitations. The outcomes of the analysis are presented in corresponding sections of the Summary report (Appendix 1). Detailed analysis of regulatory environments in the Data Submitters jurisdictions are attached as country reports.

The findings of this report should be treated as starting point and legal input for development of the framework of BIGPICTURE data repositories. The work done within the first 9 months of the project will be complemented by further efforts of WP5, in particular T5.2.1, with the aim of investigating and developing the legal and contractual framework for sustainable BIGPICTURE repository covering (nearly) all EEA countries and Switzerland.

Methods

The Summary report provides an overview on the responses collected from the country experts from the relevant jurisdictions.

In order to prepare this report, the first task was the identification of the initial Data Submitters' jurisdictions, as those were not defined in the DoA. Thus, the list of the jurisdictions involved had to be determined based on the following list of initial Data Submitters provided by WP3.

Data Submitter	Abbreviation	Country	Node
Universitair Medisch Centrum Utrecht	UMC UTRECHT	Netherlands	Cancer
Stichting Katholieke Universiteit	RUMC	Netherlands	Cancer
Azienda Ospedaliera Per L'Emergenza Cannizzaro	AOEC	Italy	Cancer
Technical University of Munich	TUM	Germany	Cancer
Philipps Universitaet Marburg	UMR	Germany	Clinical trial
The Leeds Teaching Hospitals National Health Service Trust	LTHTNHS	UK	Liver
Helsingin Ja Uudenmaan Sairaanhoitopiirin Kuntayhtymä	HUS	Finland	Lung
Medizinische Universitat Graz	MUG	Austria	Lung
*DECIPHEX**	*DECIPHEX*	*Ireland*	*Non-clinical*
Medical University Vienna	MUW	Austria	Renal
Ostergotlands Lans Landsting	RO	Sweden	Skin
*Linköping University**	*LiU*	*Sweden*	*Skin*

Universite de Liege	ULIEGE	Belgium	Skin

After further investigation, Ireland was excluded from the list based on the subsequent clarification that DECIPHEX will not be in the role of a Data Submitter. Similarly, LiU will not be participating in the role of a Data Submitter, at least during the initial period, and neither is LiU part of the skin node. For this reason, the LiU legal unit has not provided input to this document.[1] Thus the list of initial Data Submitters jurisdictions comprised the following countries: Austria, Belgium, Finland, Germany, Italy, the Netherlands, Sweden and the UK.

In parallel, TIME.LEX began mapping the areas of law which should be examined and discussing with other contributors (LYGATURE and BBMRI-ERIC) involvement of country experts. TIME.LEX developed a questionnaire for the local experts which was consulted within the task group. To ensure consistency in the level of detail and the content of the responses, TIME.LEX together with input from RO completed a report for Sweden. Next, the questionnaire and model responses were sent to identified country experts for their input.

As task leader TIME.LEX prepared an analysis of EU data protection law (General Data Protection Regulation, GDPR) and EU intellectual property law (copyright and database protection rights) and drafted the main body of the study based on own legal research and previous assessments, such as AEGLE project[2], EC reports[3] and available literature. In addition, inputs from relevant experts and partners were collected and an assessment of overarching themes and requirements was prepared to the extent they are relevant to the BIGPICTURE repository as a whole.

The country reports were prepared for Data Submitters jurisdictions in which a legal expert/respondent was available. This included respondents from the BBMRI-ERIC ELSI network covering: Italy, Finland and Belgium (jointly with the TIME.LEX team). Where no ELSI contact was available, partners from the relevant jurisdictions provided their contribution: Sweden (RO), UK (I-HD), the Netherlands (LYGATURE) and Germany (UMR).

<u>Limitations applicable to the report:</u>

- The report focuses on legal rules regarding the provision by the Data Submitters of WSIs and Metadata for BIGPICTURE project. As BIGPICTURE will deal with WSIs (digital slides), the report does not cover legal conditions regarding handling the glass slides themselves, including obtaining or transferring such original slides with Specimen.

- The report is not meant as an exhaustive publication regarding all legal issues connected with the project. The focus has been put on the legal areas outlined above. General legal considerations (e.g., on the set up of entities involved in the process, employment, general compliance rules etc.) were not examined in this report.

- Many of the discussions regarding the set-up of the BIGPICTURE project are ongoing. Thus in many cases the precise answers on legal constrains depend on outcome of these discussions and will need to be further worked out in the process of development of the project.

- Generally, the aim of the document was to assist the Data Submitters in providing their WSIs and Metadata to the project. Some parts of the report (in particular relating to data protection) also provide additional insight into internal requirements of the BIGPICTURE infrastructure, which will in turn enable the Data Submitters to take their own decisions and meet their own requirements (e.g., demonstrate a sufficient degree of data protection).

[1] Hence, both DECIPHEX and LiU have been marked with a "*" in the table above.
[2] http://www.aegle-uhealth.eu/en/
[3] For example "Assessment of the EU Member States' rules on health data in the light of GDPR" (https://ec.europa.eu/health/sites/default/files/ehealth/docs/ms_rules_health-data_en.pdf)

- The summary report is meant to provide high level comments on overarching themes and conclusions, with illustrative examples for relevant jurisdictions. Please refer to detailed responses in the country reports for more specific guidance. In the avoidance of doubt, those reports (Appendix 2-8) should be viewed as taking precedence over the Summary report (Appendix 1).

- The report does not constitute legal advice or binding interpretation of any of the discussed laws. Data Submitters are encouraged to consult their own legal teams before taking part in BIGPICTURE and making their decisions regarding sharing of the Data.

Results

Based on the reports on legal assessment performed so far, following recommendations and directions for further work can be made to the BIGPICTURE consortium:

- BIGPICTURE will most likely not be able to provide tools or guidance which will provide assurance for the anonymity of Data ("Data" meaning WSI and related Metadata jointly) for all jurisdictions in scope. These tools may be provided as means of assistance to the Data Submitters. However, the assessment whether the goal in terms of sufficiently anonymized or pseudonymized Data was reached will be always on the Data Submitter (as controller).
- Notwithstanding the above, BIGPICTURE repository should be designed to be able to handle personal data in line with the GDPR requirements, in particular pseudonymized Data.
- Privacy notices should be drafted and made available to the data subjects (including via Data Submitters), unless a specific exemption applies. This includes preparing privacy information about processing of data in the context of BIGPICTURE which would be available on the platform.
- The BIGPICTURE consortium must agree on how the process of handling Data Subject Rights ("**DSR**") will be handled internally and how the BIGPICTURE infrastructure may technically enable the exercise of the DSR rights, where legally required.
- Discussions regarding the Metadata accompanying the WSI must take into account the rule of data minimization, in particular in alignment with the research objectives for the creation of the repository. The motivation for the scope of the data should be documented.
- BIGPICTURE must take and document steps to ensure that WSIs are of representative quality (as agreed within the consortium and required for the AI training) and contain accurate personal data in the accompanying Metadata. There should also be a process of removing and correcting incorrect or misleading Data, in particular at the legitimate request of data subjects (unless exception applies).
- BIGPICTURE data controllers need to define a storage period for the pseudonymized WSI.
- Localization of the controllers of data in certain jurisdictions may trigger the applicability of their biobank laws to the BIGPICTURE collection of WSIs. This should be considered in the context of sustainability work.
- Designing and implementing appropriate organizational and technical measures to protect the personal data contained in the WSIs should be a central consideration for the project. BIGPICTURE should make the description of the safeguards available to Data Submitters and be mindful of the requirements which are provided in local legislation, to ensure that Data Submitters from those countries can participate. Some jurisdictions have concrete expectations on the measures which should be applied, while others refer to general requirements of the GDPR.
- The partners responsible for hosting the repository should be required to confirm that they appointed a data protection officer ("**DPO**") and if not - that they are exempted from this requirement. Details of DPOs should be exchanged and available for the controllers and processors of BIGPICTURE data.
- Roles of the partners (controller, joint controller and processor, no role under GDPR) need to be well understood and defined, with appropriate rights and obligations allocated via data

sharing agreements, including data processing agreements, joint controller arrangements or controller-to-controller agreements, as appropriate. These roles will need to be reflected in other privacy documentation of the project, including (but not limited to) data protection impact assessment ("**DPIA**"), privacy notices, data registers and notifications.

- BIGPICTURE should conduct a DPIA exercise to evaluate all relevant aspects of sharing of WSIs within the project, in line with the GDPR requirements. This DPIA should become a reference source that will feed into and contribute to own assessments conducted by the Data Submitters (acting as controllers).
- BIGPICTURE should carefully consider whether to mandate the Data Submitters to provide WSIs which contain Metadata on specific categories of health data (genetic data, rare diseases, Covid-19). Provision of this Metadata should be marked as optional, which may be provided that the Data Submitter may share this information in line with the national law.
  Data containing such information may need additional layer of ethical approval or special consideration.
- BIGPICTURE must engage in discussions with concrete Data Submitters on the approach of their ethical committees to delegating to Data Access Committee ("**DAC**") authorization of access to WSIs and Metadata.
- BIGPICTURE License(s) should be drafted to contain license (or transfer of rights) for the use of intellectual property rights to the submitted WSIs for the purpose of the BIGPICTURE project. License requirements should also be included in the contract (or terms and conditions) of the use of Data by researchers accessing the repository.

Moreover, the initial list of responsibilities which should be addressed by the Data Submitters includes the following:

- Data Submitters should indicate whether they are submitting pseudonymized or anonymized Data. For pseudonymized Data, the Biological Being ID mapping key (which allows the Data to be re-identified) must be kept separately, in particular it should not be provided to BIGPICTURE repository. Moreover, the key must be protected by technical and organisational measures (e.g. on the premises of Data Submitter).
- If the Data Submitter is contributing anonymized data, it should evaluate that the patient would not be identifiable from the submitted Metadata or the WSI itself. In particular, the Data Submitter must confirm that it performed an assessment that with reasonable effort the patient cannot be re-identified from the Data (e.g., due to number of different data points in the Metadata or their unique combination), despite certain patient identifiers (such as name or ID number) being removed and despite the assigned Biological Being ID being a "random" value which is not mapped against the actual patient record by the Data Submitter.
- Data Submitter assures that it has the appropriate legal basis to contribute the Data (WSIs and accompanying Metadata) to BIGPICTURE. The applicable legal basis must comply with the GDPR and with Data Submitter's national law and practice on re-use of health data for research purposes and take into consideration that the submitted Data may be made available to other researchers via BIGPICTURE repository in line with the project's objectives.
- If Data Submitter obtained consent from data subjects, Data Submitter assures such consent fulfils the requirements of the GDPR for explicit consent.
- Data Submitter assures that the submission of any Data derived from clinical trials is in line with any consent obtained from the individual and/or information provided to them during the clinical trial.
- Data Submitter confirms that the Data that it makes available to BIGPICTURE do not contravene any of their national laws restricting the disclosure of specific types of personal or health information, such as national ID number or information about certain diseases.
- Data Submitter must verify if data subjects need to be informed about the contribution of their Data and if yes, confirms informing them about it.
- Data Submitter verifies whether they are obligated to ensure that data subjects may exercise their DSR and informs BIGPICURE about it.

- Data Submitter checks whether the safeguards offered by BIGPICTURE fulfil their requirements.
- Data Submitter confirms that they appointed a DPO and provides DPO's contact details. If no DPO was appointed, Data Submitter must confirm that they are exempted from this requirement.
- Data Submitter confirms that they obtained any required ethical approval for participating in BIGPICTURE and that submission of Data and making WSI and Metadata available to researchers via the repository is allowed in line with the conditions of such approval.
- Data Submitter confirms whether the WSI and related Metadata are submitted in pseudonymized or anonymized form, as defined by laws, regulations and organizational policies that govern the Data Submitter, and that the data have been de-identified in line with any requirements or guidance that BIGPICTURE may produce.
- Data Submitter confirms that they have removed any non-personal information from WSI and Metadata to the extent required to protect the IP rights or other proprietary information of the Data Submitter (e.g., constituting trade secrets of the Data Submitter or required by the national law).
- Data Submitter states that it is not subject to any special rules under national biobank law which would restrict the use of WSIs and Metadata in BIGPICTURE. Data Submitter indicates whether it is under a requirement for concluding material transfer agreement (MTA).
- Data Submitter confirms that they have cleared any third-party IP rights to Data, in particular that the Data Submitter has obtained such rights, either by virtue of law or by contract (e.g., by an agreement with the researcher).
- Data Submitter confirms that they are not subject to open-data laws which would constitute an obstacle in providing the WSIs to the project nor that they impose any conditions for Data use in the context of BIGPICTURE.
- Data Submitter agrees to conditions of the data transfer agreement which may be developed by the project, including to conditions of processing of the submitted personal data (if the WSI have been pseudonymized) and the BIGPICTURE License terms for the use of WSI which are agreed on during submission.

The lists above are to be considered as preliminary, to be further developed in more granularity as the project progresses.

Completed country reports summarizing legal frameworks in Belgium, Finland, Germany, Italy, the Netherlands, Sweden and UK are provided in Appendix 2-8 to this document. Based on the country reports, a Summary report was prepared and is attached as Appendix 1. Summary report follows the outline as indicated in the table of content above.

Discussion

The title of the deliverable (D5.1: Overview of the legal frameworks of all countries involved in BIGPICTURE) is misleading. Further to the elaboration of the work in T5.1, the object of the report was to provide a summary the rules *applicable to the Data Submitters initially involved in the project* and not all the countries involved in the project. As a result of prioritization of the countries to be examined and availability of legal experts, detailed assessments of 7 (seven) Data Submitters' jurisdictions were made. This initial report will be extended during the work of WP5 to cover the regulatory frameworks of (nearly) all EEA countries plus Switzerland.

Conclusion

The summary report clearly demonstrates that many legal requirements for the BIGPICTURE project stem from the EU and applicable national rules on processing and re-use of health data for scientific purposes. These obligations are fragmented and vary to a great extent, depending on the country of the Data Submitter, type of the Data Submitter and the context of collection of the initial slide and

sometimes even on the local interpretation of the law. While the Data Submitters may be advised on the existence of the laws, the decision on what is the appropriate manner to address them must be left to the Data Submitters as the initial controllers and data holders.

Moreover, the ethical requirements for submitting personal data contained in the pseudonymized clinical WSI to publicly accessible repositories are not well defined by law and often depend on the decision of the ethics committee. There are very limited rules relating specifically to the use of WSIs and related Metadata and thus the legal principles have to be deducted from often very general laws and provisions. This results in vague and often conflicting requirements stemming from various legal acts. Determining the applicable details will in our view inevitably entail engaging in discussions with particular Data Submitters and “one model fits all” solutions will be difficult to find.

Despite these variances, some legal aspects, even though regulated by different national laws, can be summarized into overarching themes. This results in the recommendations provided in Results section above. They will be further developed and implemented to BIGPICTURE throughout the work done by WP5.

Attachment list

Appendix 1: Summary Report
Appendix 2: Belgium Country Report
Appendix 3: Finland Country Report
Appendix 4: Germany Country Report
Appendix 5: Italy Country Report
Appendix 6: Netherlands Country Report
Appendix 7: Sweden Country Report
Appendix 8: UK Country Report

Appendix 1

Summary report on overview of the legal frameworks of the Data Submitters initially involved in the BIGPICTURE project

Glossary and Abbreviations

The definitions and abbreviations provided below relate to this deliverable only. The terms marked with "*" have been aligned with definitions of *D1.04 - Initial Data Management Plan*.

"**Biological Being**"* represents a patient or an animal from whom/which specimen are acquired to produce slides or associated data (genetics, laboratory measurements, clinical history, etc.).

"**Whole Slide Image (WSI)**"* refers to the scanned copy of a glass slides. For the purpose of the deliverable, WSI are split into "clinical WSIs" of human origin Specimen and "non-clinical" WSIs of animal Specimen.

"**Metadata**" means "data about data." In the context of this document it refers to any and all data linked to the WSI, consisting of data relevant to the project such as technical data, license, ethical and legal compliance data, biological data (e.g., target, organ, animal, study id…), compound data (e.g., compound ID, structure, …), slide annotation data, and other associated data (e.g., study reports and protocols, clinical data, clinical pathology and biomarker data, macroscopic examination report, …).

"**Data**" means WSI and Metadata jointly.

"**Data Access Committee (DAC)**" means organisation, individual or a group of individuals which is authorized to decide on the access requests to the contributed Data.

"**Data Subject Rights**" mean the rights referred to in Articles 15-21 General Data Protection Regulation.

"**Data Submitter**" refers to the party (contributor) providing the Data by uploading WSI and Metadata to the repository.

"**General Data Protection Regulation**" means 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

"**Specimen**"* refers to biological material that is being extracted from the person/animal (Biological Being). The term does not refer to the container in which those parts are placed, and sometimes parts from different anatomical locations are placed together in such a container, making the origin of a given Specimen ambiguous. Hence, Specimen must always only be derived from one specific anatomical location.

"**Slide**"* refers to the microscope (glass) slide as a physical object from which a WSI is generated.

"**BIGPICTURE License**"* refers to the legal arrangement (terms of use) between the Data Submitter and the end-user, or the place the Data will be deposited, specifying what users can do with the Data defined at Data submission. There are likely to be several types of BIGPICTURE Licenses to be chosen from by the Data Submitter.

The terms defined in the GDPR, such as "**personal data**", "**controller**", "**processor**", "**data subject**", "**pseudonymization**" will be used in the meaning of this act, however will not be capitalized in this document.

AI	Artificial Intelligence
DAC	Data Access Committee
DoA	Description of Action under GA945358
DPIA	Data Protection Impact Assessment

DPO	Data Protection Officer
DSR	Data Subject Rights
EDPB	European Data Protection Board
GDPR	General Data Protection Regulation
ISO	International Organisation for Standardisation
WP	Work Package
WSI	Whole Slide Image

Background

Based on the DoA and reconfirmed thorough discussions with WP3, the following facts and assumptions were taken into considerations when preparing the report[4]:

- The clinical part of the BIGPICTURE data repository will consist of digital scans of slides of human tissues or biological samples (WSI, whole slide images). The original slides were collected either for clinical care purposes or in the context of clinical studies. Thus, the original slides will not be collected directly and primarily for the purpose of digitization for the BIGPICTURE data repository.
- The BIGPICTURE data repository will also include WSI from animal tissues (non-clinical WSIs) that are also going to be digitized slides from existing collections. The BIGPICTURE repository will only include digital slides (WSIs), not the glass originals or biological samples.
- The WSI will be accompanied by Metadata (i.e. not only image will be stored). Scope of the metadata may vary, depending on the type of WSI. Metadata parameters will be defined by the Metadata & Nomenclature Task Force.
- The purpose of BIGPICTURE is to store WSI and related Metadata ("**Data**") and facilitate access to Data for scientific and research purposes.
- The Data Submitters will have or will digitize Slides generated as part of clinical diagnostic work that, in some instances, will also have been used in clinical studies. Some Data Submitters may also contribute medical data which was generated as part of the research projects which involved the collected samples. There may be additionally data from archives[5].
- No animal experimentation will be performed for the purpose of data collection in the project.
- Additional patient information may accompany WSI in the Metadata. The final list will be provided at a later stage, as outlined in D3.2. Results from genomic investigation may be used as associated data in some WSI collections.

1. Protection of personal data

Initial Data Submitters are located in the Member States of the European Union (EU), with the exception of one submitter (i.e. LTHTNHS) located in the United Kingdom (UK). Therefore, in this part of the report we will discuss the rules resulting from the General Data Protection Regulation (EU) 2016/67910 ("**GDPR"**)[6] and the national laws applicable to protection of personal data.

The GDPR provides the basic provisions constituting the general legal framework applicable in all Member States of the European Economic Area in relation to the protection of personal data. In the

[4] Please note that these assumptions have been updated to reflect the project's progress, when compared to the background information provided initially to the country experts (as seen in the country reports). The updates and purpose of BIGPICTURE have been however discussed with the country experts during collection and refining of the reports.
[5] E.g. results from blood tests, diagnoses, treatment, etc. that are kept as patient records in the hospital information system and/or in a Research database (RDA, https://www.meduniwien.ac.at/web/mitarbeiterinnen/it-hilfe-support/it4science/plattformen/rda/).
[6] 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) (OJ L 119, 4.5.2016, p. 1–88).

UK, the GDPR is retained in domestic law as the "UK GDPR." For now, the rules of the (EU) GDPR discussed below are also in place in the UK under the UK GDPR. However, for the future, UK will have the independence to revise its framework.

When considering the legal frameworks relating to protection of data in the EU and UK, also national laws of the Data Submitters jurisdictions need to be examined. As observed by the study "Assessment of the EU Member States' rules on health data in the light of GDPR" conducted recently by the European Commission in the context of preliminary preparations for the creation of EHDS ("**EC Health Data Study"),** the laws are very fragmented and vary from country to country. The authors of the study state: "*The GDPR provides that Member State legislators may adopt legislation to allow for use of data for research in accordance with Article 9(2)(j) and 89(1)). It is clear from the responses provided by the correspondents that the Member States have not implemented such legislation in a homogenous way,* ***resulting in a complex and fragmented landscape for researchers to navigate.*** *Consequently, differences between Member States in the way the GDPR is implemented and interpreted in the area of scientific research has made data exchange between Member State and EU bodies for research purposes difficult and in some cases highly technical."*[7] Furthermore, according to the combined evaluation roadmap/Inception Impact Assessment[8] for the EDHS initiative while the GDPR sets out the EU data protection rules, the "*Member States may* ***further specify some aspects in specific areas, such as health data, which they have done to a large extent****. As a result, the processing of personal health data in Member States is fragmented, leading to obstacles and to limited access of researchers and public institutions, that in turn reduces the EU competitiveness and innovation potential at a global level*."

Being aware of the varying national rules on the (re)use of health data, a number of questions (Q1-Q7) in the questionnaire for the country respondents focused on those rules. Summary of the findings from the received responses will be provided below. However, in order to discuss those rules, it is pivotal to assess whether and to what extent the WSI and related Metadata should be considered as "personal data" under GDPR, thereby triggering its applicability. After this discussion, we will outline the principles of the data protection, which will need to be considered in the project, with the focus on the specific rules of national laws which are applicable to BIGPICTURE Data Submitters.

1.1. Definition of "personal data" in the context of WSI

The term "**personal data**" is broadly defined in the GDPR as any information relating to an identified or identifiable natural person (data subject)".[9] An identified or identifiable natural person for the purposes of the GDPR is "one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person."[10] Thus, the term personal data comprises not only directly identifiable information (such as names, addresses, contact information, national ID number), but also indirectly identifiable information (such as pseudonymous information where data can only be linked to a specific semantically meaningless number, such as alphanumeric code).

The information must make that person "**identified** or **identifiable**." An **identified person** is distinguishable from other members in a group. A person described by a data record is considered to be **identifiable** if any actor exists at present or in the future who is able to identify the data subject by using any realistically available additional information and linking methodology.[11] Identification can be direct – meaning that the initial data record and information available to the actor are linked without the use of additional information - or indirect – meaning that initial data record needs to be linked with

[7] Assessment of the EU Member States' rules on health data in the light of GDPR (https://ec.europa.eu/health/sites/default/files/ehealth/docs/ms_rules_health-data_en.pdf), further referred to as "EC Health Data Study", pg. 58.
[8] https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12663-Digital-health-data-and-services-the-European-health-data-space_en
[9] Article 4(1) GDPR.
[10] *Ibidem.*
[11] This definition and further analysis of concepts of identified or identifiable person in: "Towards a Better Understanding of Identification, Pseudonymization, and Anonymization" (https://www.datenschutzzentrum.de/artikel/1372-.html).

additional information and then with information available to the actor[12], thus multiple pieces of information must be put together.

In this context it is important to delineate two terms, which relate to the definition of personal data: pseudonymization and anonymization.

Article 4(5) GDPR defines **pseudonymization** as "the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organizational measures to ensure that the personal data are not attributed to an identified or identifiable natural person."

"**Anonymous information**" is one which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable (recital 26 GDPR). When putting together information to make a person identifiable, recital 26 to the GDPR provides that the threshold is that it must concern "means reasonably likely to be used." In assessing such *reasonableness*, "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." Important to note is that this notion of *reasonableness* can evolve over time. It may occur that at the start of a processing operation the state of technology does not make identifiability of a person reasonably likely. However, if the processed information is stored for a certain duration, technology may become available during that time that does make identification reasonably likely. Given the rapid development of data mining tools, it is therefore very likely that data previously not considered as personal data may sooner or later become personal data. This means that any kind of information should be assessed on a case-by-case basis to ascertain whether it renders a natural person identified or reasonably likely to be identifiable.

Based on the information received from the country reports, human WSI with patient Metadata will be considered as personal data if this Metadata will allow linking the WSI back to the patient. In particular, the **German** report notes that: "*it should be noted that WSI represent high-resolution images of a tumor. The use of tumor images for research and education has not been restricted by data protection regulations(…) it is not possible to identify the patient from an image of his/her tumor, therefore images represent anonymous information. If the tumor image is linked to other biological datasets of an individual (e.g., via metadata which are part of additional datasets), more complex regulation[13] are relevant."* Also, in **Belgium**: "*Associated data, more specifically personal data related to health, are collected alongside samples, hence, in the context of HBM [HBM meaning "human body material" - comment by TLX] procurement, the data protection rules apply too. Images of human body samples are "data" ("personal data" if the person concerned is identifiable)."*

In BIGPICTURE, as described in *D3.2 Minimal dataset deliverable*, each WSI's Metadata (either from human or an animal) will contain a "Biological Being ID" which is "a unique identifier of the Biological Being", for example "PAT0001." While this is not the deciding criterium (see further comments below), yet for human WSIs, if the controller (Data Submitter) or any other party will be able to trace back the Biological Being ID to the individual from which the Specimen was taken, this would indicate that the WSI is "identifiable" information. This would in particular be the case, if the Data Submitter keeps a key mapping of the Biological Being IDs with the patients' identities at its premises.[14]

Furthermore, if the WSI will contain also Metadata about the diagnosis of the person (defined in D3.2 as "diagnosis that is associated to the entire image", either in English or based on established coding system), WSI would contain "data concerning health" which is a special category of personal data

[12] "Towards a Better Understanding…", pg. 25.
[13] The said regulations are consent rules which are applicable to personal data.
[14] At the time of preparing of this report, we are aware that some of the Data Submitters will be mapping the IDs and then providing anonymous data. For instance, RÖ plans to submit anonymous data only, with no possibility to link submitted data back to patient identities, which is indeed all that the ethical review approval that they obtained for this activity allows. Biological Being ID will be randomly assigned, and only guaranteed to be consistent inside each dataset or study.

under GDPR[15]. These data would be pseudonymized, assuming that key which maps the Biological Being ID codes with the patient data will be kept separately by the Data Submitter and protected by technical and organisational measures and that the remaining Metadata or WSI itself would not allow to identify the patient. This is because data are not pseudonymous if a specific person is identifiable from the data solely without additional information. This could happen when indirect identifiers and exceptional records enable identification, even if direct identifiers (e.g., names) are stored separately and securely. Pseudonymisation is also unsuccessful if an outside person is able to determine the original values based on the pseudonyms.

Even when the Data Submitter assigns the Biological Being ID randomly and does not keep a list of the IDs mapped against the identity of the patient, this may not by itself be sufficient to state that the Data is anonymous. This because Data are anonymous and fall outside of the scope of the GDPR only if - considering all the circumstances[16] - the threshold of "identifiable" information is not met. This threshold is particularly difficult to reach for health data, while retaining its usefulness for the research which is to be conducted, not to mention complying with arising obligations of using quality data in research (e.g., for AI training[17]). In particular, in its guidance on Ethics & Governance of Artificial Intelligence for Health[18], World Health Organization points out that "*Anonymization may not be possible during health data collection. For example, in predictive AI, time-course data must be collected from a single individual at several times, obviating anonymization until data at all time points are collected. Furthermore,* ***while anonymization may minimize the risks of (re-)identification of a person, it can reduce the positive benefits of health data****, including re-assembly of fragments of an individual's health data into a comprehensive profile of a patient, which is required for some forms of AI such as predictive algorithms of mortality. Furthermore, anonymization may undermine a person's right to control their own data and how it may be used. (…)"*

Also, a paper published by the Spanish Data Protection Agency (AEPD) and European Data Protection Supervisor (EDPS) on "10 misunderstandings related to anonymisation" states that it is a common misunderstanding that anonymization of data is always possible. Moreover, those authorities observe that "*it is not always possible to lower the re-identification risk below a previously defined threshold whilst retaining a useful dataset for a specific processing*."[19]

In this context, an important note is to be found in Article 29 Data Protection Working Party (WP29)[20] in its Opinion 05/2014 on Anonymization Techniques (WP216). WP29 stated that "*when a data controller does not delete the original (identifiable) data at event-level, and the data controller hands over part of this dataset (for example after removal or masking of identifiable data), the resulting dataset is still personal data*"[21]. Following this interpretation, if the decision was taken on anonymization of data, the Data Submitter would need to delete the medical records containing WSI from their systems[22]. Moreover, also in its opinion the Article 29 Working Party admits that truly anonymous information is rare. Many so-called anonymization techniques can be reversed, or still result in information that could be coupled with other information to make a natural personal reasonably identifiable.

[15] Under Article 4(15) GDPR '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. Article 9 GDPR provides additional rules for processing of special categories of personal data, including data concerning health.
[16] Some practical examples can be found in https://www.fsd.tuni.fi/en/services/data-management-guidelines/anonymisation-and-identifiers/.
[17] According to 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 (COM/2021/206 final) available here: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206, data sets used to train high-risk AI systems will need to fulfil quality criteria described in Article 10 (2)-(5), which include, inter alia, that they must be relevant, representative, free of errors and complete, and have appropriate statistical properties.
[18] https://www.who.int/publications/i/item/9789240029200
[19] https://edps.europa.eu/data-protection/our-work/publications/papers/aepd-edps-joint-paper-10-misunderstandings-related_en
[20] Article 29 Data Protection Working Party is now replaced by the European Data Protection Board, however the position papers of WP29 are still often referred to, even without formal endorsement decision.
[21] https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf
[22] Please note that this WP29 interpretation is not legally binding and is often contested by researchers as the correct approach (e.g., Khaled El Emam, Cecilia Álvarez, A critical appraisal of the Article 29 Working Party Opinion 05/2014 on data anonymization techniques, International Data Privacy Law, Volume 5, Issue 1, February 2015, Pages 73–87, https://doi.org/10.1093/idpl/ipu033). Moreover existing repositories of medical images often claim that the data is provided as "anonymous".

In the context of BIGPICTURE, the responses from the legal experts confirm that **there is neither universal understanding or threshold for anonymization of health data nor conclusive guidelines issued by relevant authorities.** For instance, as noted by the **UK** report "*generally speaking, whether the data should be considered anonymous is evaluated from the perspective of the recipient of the data. Ultimately, it is the hospital which decides whether the data is anonymous or not (…).*" While the **Finnish** expert notes that "*The level of de-identification of data that would fulfil the criteria of anonymous data further to understanding adopted by the Finnish data protection authorities (in line with the WP29 opinion on data anonymization) may be difficult to achieve in relation to health data, in particular WSI.*"

This factor is likely to impact the BIGPICTURE project. Moreover, if the human WSI will include a Metadata identifier which will allow any party (including the Data Submitter) to link the WSI to the individual patient (for instance because the Data Submitter will keep a key which links the assigned Biological Being IDs with patient names), they would not be viewed as anonymous in the light of the GDPR. However, as noted above, even lack of linking of the Data to the patient via mapped Biological Being ID does not necessarily mean that the Data is anonymous.

All in all, given the aims of the project, the repository should be built to have the capability to process sensitive personal data contained in the contributed WSIs and Metadata, however this does not imply that all WSIs will be personal data, as some Data Submitters may chose to provide anonymous information.

Conclusion: **BIGPICTURE will most likely not be able to provide tools or guidance which will provide assurance for the anonymity of slides for all jurisdictions in scope. These tools may be provided as means of assistance to the Data Submitters. However, the assessment whether the goal in terms of sufficiently anonymized or pseudonymized data was reached will always be on the Data Submitter (as the controller).**

Conclusion (2): **If the Data Submitter is contributing pseudonymized data, the Biological Being ID mapping key must be kept separately by the Data Submitter (in particular the key should not submitted to BIGPICTURE) and protected by technical and organisational measures (e.g., secured at the Data Submitter's premises). It must be checked that the disclosed Metadata does not allow patient identification even without access to the key.**

Conclusion (3): **If the Data Submitter is contributing anonymized data, they should verify that the patient would not be identifiable from the Data, even if Biological Being ID is a "random" value which is not mapped against the actual patient record. In particular, Data Submitter must evaluate that the patient cannot be identified due to number of different data points in the Metadata or their unique combination, despite certain identifiers (such as name or ID number) being removed.**

Please note that further comments in section 1 below are provided based on the assumption that (some of) the clinical WSIs will be provided as pseudonymized data and thus GDPR will apply to them.

1.2. GDPR key requirements in the context of BIGPICTURE

The GDPR implies compliance with seven key principles:[23]

- lawfulness, fairness and transparency – meaning that a legal basis for any data processing (including but not limited to consent) must be available, and that the persons concerned must be appropriately informed of how their data will be used;
- purpose limitation – meaning that data must be collected for specific purposes, and may thereafter only be used for compatible purposes;

[23] As enshrined in Article 5 (1) and (2) GDPR.

- data minimisation – meaning that data collected and used in the project must be as minimal as possible, taking into account the intended purposes;
- accuracy – meaning that measures must be taken to ensure the quality and accuracy of the data, and that measures must be available to detect and remedy problems;
- storage limitation – meaning that data may only be retained for as long as necessary given the intended purposes, and that it must thereafter be deleted or anonymised;
- integrity and confidentiality – meaning that data must be protected by appropriate technical and organisational measures to ensure its confidentiality, integrity and availability;
- accountability – meaning that responsible entities must be identified, and that appropriate controls (such as logs) are available to ensure that any problems can be attributed to the correct entity.

While a detailed discussion of all the principles is beyond the scope of this report, below we discuss the key considerations for BIGPICTURE Data contribution and repository set up, as provided in the GDPR and country expert reports. The report does not aim to provide an exhaustive list of all potential privacy issues which need to be dealt with within this project. For instance, given the scope of the project, other matters which need to be addressed and accounted for should include:

- transfers of personal data outside of EEA (in compliance with Chapter V GDPR),
- detection and reporting of data breaches (in compliance with Articles 33 and 34 GDPR),
- consultations with supervisory authorities (in particular, in circumstances foreseen under Article 36 GDPR).

1.2.1. Lawfulness

Under the GDPR the concept of "processing" is very broadly defined[24] and thus submission of the clinical WSIs containing personal information (as noted above) to the BIGPICTURE project would be considered "processing of personal data". Thus, the main question for Data Submitters (acting as controllers) contributing the clinical WSIs is determining whether they can substantiate a legal basis for the WSIs and Metadata to be submitted and used.

Articles 6 and 9 GDPR provide the following legal basis for processing of personal data.

Article 6.1
1. Processing shall be lawful only if and to the extent that at least one of the following applies: (a) the data subject has given consent to the processing of his or her personal data for one or more specific purposes; (b) processing is necessary for the performance of a contract to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract; (c) processing is necessary for compliance with a legal obligation to which the controller is subject; (d) processing is necessary in order to protect the vital interests of the data subject or of another natural person; (e) processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller;

[24] Article 4(2) GDPR 'processing' means any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction;

(f) processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child. Point (f) of the first subparagraph shall not apply to processing carried out by public authorities in the performance of their tasks.
Article 9
1. Processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation shall be prohibited. 2. Paragraph 1 shall not apply if one of the following applies: (a)the data subject has given explicit consent to the processing of those personal data for one or more specified purposes, except where Union or Member State law provide that the prohibition referred to in paragraph 1 may not be lifted by the data subject; (b)processing is necessary for the purposes of carrying out the obligations and exercising specific rights of the controller or of the data subject in the field of employment and social security and social protection law in so far as it is authorised by Union or Member State law or a collective agreement pursuant to Member State law providing for appropriate safeguards for the fundamental rights and the interests of the data subject; (c)processing is necessary to protect the vital interests of the data subject or of another natural person where the data subject is physically or legally incapable of giving consent; (d)processing is carried out in the course of its legitimate activities with appropriate safeguards by a foundation, association or any other not-for-profit body with a political, philosophical, religious or trade union aim and on condition that the processing relates solely to the members or to former members of the body or to persons who have regular contact with it in connection with its purposes and that the personal data are not disclosed outside that body without the consent of the data subjects; (e)processing relates to personal data which are manifestly made public by the data subject; (f) processing is necessary for the establishment, exercise or defence of legal claims or whenever courts are acting in their judicial capacity; (g)processing is necessary for reasons of substantial public interest, on the basis of Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject; (h)processing is necessary for the purposes of preventive or occupational medicine, for the assessment of the working capacity of the employee, medical diagnosis, the provision of health or social care or treatment or the management of health or social care systems and services on the basis of Union or Member State law or pursuant to contract with a health professional and subject to the conditions and safeguards referred to in paragraph 3; (i) processing is necessary for reasons of public interest in the area of public health, such as protecting against serious cross-border threats to health or ensuring high standards of quality and safety of health care and of medicinal products or medical devices, on the basis of Union or Member State law which provides for suitable and specific measures to safeguard the rights and freedoms of the data subject, in particular professional secrecy;

(j) processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) based on Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject.

3. Personal data referred to in paragraph 1 may be processed for the purposes referred to in point (h) of paragraph 2 when those data are processed by or under the responsibility of a professional subject to the obligation of professional secrecy under Union or Member State law or rules established by national competent bodies or by another person also subject to an obligation of secrecy under Union or Member State law or rules established by national competent bodies.

4. Member States may maintain or introduce further conditions, including limitations, with regard to the processing of genetic data, biometric data or data concerning health.

In general, as follows from desk research and the country expert reports, the landscape regarding use of the above listed legal basis for the reuse of health data collected for other purposes (in particular in clinical or clinical trial context) is very fragmented. More specifically, the permissible legal basis largely depends on national laws and practice. This conclusion is concurrent with the finding of the EC Health Data Study[25] that "*It is clear from the views shared in the workshops and by country correspondents to the legal and technical survey that while the GDPR is a much appreciated piece of legislation,* ***variation in interpretation of the law and national level legislation linked to its implementation have led to a fragmented approach which makes cross-border cooperation for care provision, healthcare system administration or research difficult***." Furthermore: "*Given that the GDPR foresees the possibility of special legislation for processing of genetic information, it is not surprising that* ***significant variation*** *may exist between some Member States in this area"*[26]. Moreover, the national laws may provide specific rules regarding use of health data (Article 9.4 GDPR).

However, while potentially different legal bases may be considered, in practice what transpires from the reports is that the "choice" of the legal basis of use for the Data Submitter would be narrowed to the following options:

- obtaining consent from the patients as data subjects (Article 6(1)a and Article 9(2)(a) GDPR) or
- "scientific privilege" i.e. re-use of existing data collected for other purpose under certain conditions and safeguards (Article 5(1)(b) and 89(2) GDPR).

The latter possibility may be also conflated with a separate legal basis for re-use of data and considered provided for under Article 6(1)(e) or (f) in conjunction with Article 9(2)(j) GDPR and, if applicable, national provisions. Such approach seems to be presented by Article 29 Working Party Guidelines on consent under Regulation 2016/679 which explicitly point out that "*The GDPR does not restrict the application of Article 6 to consent alone, with regard to processing data for research purposes. As long as appropriate safeguards are in place, such as the requirements under Article 89(1), and the processing is fair, lawful, transparent and accords with data minimisation standards and individual rights, other lawful bases such as Article 6(1)(e) or (f) may be available. This also applies to special categories of data pursuant to the derogation of Article 9(2)(j)*"[27]. This is further discussed in "Purpose limitation" section, while below we will focus on consent as one of the possible options for submitters.

In the context of BIGPICTURE, consent was indicated as the dominant basis for the submitting and use of WSI and Metadata for research purposes, with limited exceptions to this rule.

In particular, in **Italy** *"Prior informed consent of the patient is the general legal basis for research on*

[25] https://ec.europa.eu/health/sites/default/files/ehealth/docs/ms_rules_health-data_en.pdf, further as "EC Health Data Study", pg. 9, Executive summary.
[26] EC Health Data Study, pg. 26
[27] Guidelines on Consent under Regulation 2016/679 (wp259rev.01) https://ec.europa.eu/newsroom/article29/items/623051/en, pg. 28.

health data and samples. The art. 110 of the D.Lgs 196/2003 (Personal data protection code) provides exceptions to the consent rule" (response to Q3). Consequently, "*if consent cannot be obtained (art. 110 D.Lgs. 196/2003), or it requires a disproportionate effort, the research project must be approved by the local ethical committee and the research project must be submitted to prior consultation with the Italian Privacy Authority* "(response to Q32).

In **Germany** *"In most cases patient's consent is required (…) Using patient's data without consent requires a dataset-based examination (…) In this examination, the data protection concept, major interest of research and the patient's protection-worthy interest in the data are being are balanced against each other. The researcher's academic freedom must outweigh patient's fundamental right of informational self-determination. If this is the case, the data / material might be used. Specifically for WSIs, if those conditions are fulfilled, the data provider could provide the slides to the project, even if they were collected initially for another purpose"* (response to Q4).

In **Finland**, the situation is different: "*Consent is required for physically or psychologically invasive research (…) but for processing personal data consent is just one of the options for legal bases. The questions and this response largely relate to use of data not initially collected for research. Secondary use of samples/data is possible subject to the rules described (…)"* (response to Q4). Specifically, the **Finnish** law provides specific rules on the use of health data for scientific research purposes, if suitable and specific measures to safeguard the rights of the data subject are in place.

Also in **Belgium** the law "*does not impose the use of consent as a legal ground for (primary) data processing, hence the data processing may be based on any of the six legal bases described in Article 6(1) of the GDPR. Regarding further processing of data, the rules established in the GDPR and the rules of the Belgian Data Protection Law, as cited above, apply. However, informed consent as an ethical requirement for procurement of HBM for scientific research is needed, with some exceptions"* (response to Q4). Nevertheless, the consent rules still apply (with some exceptions) with respect to obtaining and using human biological material.

Given the above, it would not be possible to indicate to the Data Submitters which legal basis is appropriate for their submission of their Data, much less impose a particular basis (e.g., consent) for everyone. The expert reports indicate that ensuring a legal basis should be a responsibility of the Data Submitter and the conditions and the applicable basis vary, depending on the national law and also on the context in which the data were collected. Moreover, contingent on the set up of the roles of the partners in the consortium (see section 1.2.8.2 Roles of the involved parties below), other partners which will process the data as controllers, may need to validate whether they would be permitted to do so. The project will further investigate those rules during the definition of the honest broker mechanism (WP3) and more detailed work on the set-up of the repository.

Last, but not least, in the event that the Data Submitter decides to collect consents from the data subjects it would need to ensure that such consent meets the requirements of the GDPR.[28] Fulfilment of those conditions is not easy, in particular it may require re-consenting for use of data which were collected from patients in another context. This may lead to decrease in the availability and diversity of the data.

Conclusion: **Each Data Submitter must check whether they can assure that they have the appropriate legal basis to contribute Data to BIGPICTURE. In many cases consent of the data subject may be the appropriate legal basis, however ensuring that this consent fulfils the requirements of the GDPR is not easy.**

1.2.2. Fairness and transparency

The rule of transparency states that data must be processed in a transparent matter in relation to the data subject. This rule is closely linked to information rights (Art. 13-14 GDPR) and data subject rights

[28] Article 4 (11) provides for definition for consent under which it is freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her. Article 7 GDPR provides for further conditions of consent.

("DSR", Art. 15-21 GDPR). Under those provisions, in general, data subjects must be properly informed about the details of processing of their data for a specific purpose (also in case the data are not collected directly from the data subject).[29] There are very limited exceptions from this requirement. For example, under Article 14(5)(b) GDPR, if the data are collected **not directly** from the data subject, the requirement to provide him/her with information does not apply if the provision of such information proves impossible or would involve a disproportionate effort, in particular for processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, subject to the conditions and safeguards referred to in Article 89(1) or in so far as the obligation to provide information is likely to render impossible or seriously impair the achievement of the objectives of that processing.[30] If the data are collected **directly** from the data subject, this exception does not apply.

In the context of BIGPICTURE, the transparency of the processing vis-à-vis the data subjects should be ensured through appropriate notices and policies, such as properly drafted information language made available to patients (unless exception applies), privacy notice on the repository website and clear description in any agreements to be entered into by the Data Submitters, such as BIGPICTURE License.

Data subject rights give the data subject possibilities to obtain information about processing and - in cases specified by the law - influence such processing.

These rights are set out in Chapter III of the GDPR, and include the rights of the data subject to :

- obtain information about the processing of their personal data (Article 15(1) GDPR)
- obtain access to the personal data held about them and copy of the data (Article 15(1) and 15(3) GDPR)
- ask for incorrect, inaccurate or incomplete personal data to be corrected (Article 16 GDPR);
- request that personal data be erased when it's no longer needed or if processing it is unlawful (right to be forgotten, Article 17 GDPR);
- request the restriction of the processing of their personal data in specific cases (Article 18 GDPR);
- receive personal data in a machine-readable format and send it to another controller ('data portability', Article 20 GDPR);
- object to the processing of their personal data for marketing purposes or on grounds relating to their particular situation (Article 21 GDPR);
- request that decisions based on automated processing concerning them or significantly affecting them and based on their personal data are made by natural persons, not only by computers. They also have the right in this case to express their point of view and to contest the decision (Article 22 GDPR).

However, according to Article 89(2) GDPR there are possible exceptions to the rules of transparency. In particular, where personal data are processed for scientific or historical research purposes or statistical purposes, Union or Member State law may provide for derogations from the rights referred to in Articles 15, 16, 18 and 21 GDPR subject to the conditions and safeguards referred to in paragraph 1 of this Article in so far as such rights are likely to render impossible or seriously impair the achievement of the specific purposes, and such derogations are necessary for the fulfilment of those purposes.

According to the country reports, the following countries made use of this possibility: **Belgium**, **Finland, Germany,** the **Netherlands**.

In **Belgium**, while exceptions are provided, the regime of making use of them is quite complex: "*A*

[29] Detailed list of information to be provided is included in Articles 13 and 14 GDPR.
[30] While whether the exception applies needs to be decided on a case by case basis, some arguments which may be invoked relate to consequences of providing information to the data subjects to the research objectives. In other cases, it may not be possible to contact the patients which have been treated long time ago or re-contacting them may be unethical.

controller for scientific or historical research or for statistical purposes may consider that he can comply with all the provisions of the GDPR and therefore not rely, or only partly rely, on the derogations provided for according to Art. 89(2) of the GDPR" (response to Q3). Thus, there are various scenarios which may be followed by the controller. If however the controller wishes to take benefit of the derogations, "*Title 4 provides for appropriate safeguards which the controller wishing to derogate from the rights of data subjects must implement."*

Also in other jurisdictions there are certain conditions which apply if controller wishes to makes use of the exceptions from data subjects' rights. In particular in **Finland,** while there are statutory deviations from the rights provided in Articles 15, 16, 18 or 21, it is also required that: "*1) the processing is based on an appropriate research plan; 2) a person or group responsible for the research has been designated; and 3) the personal data are used and disclosed only for scientific or historical research purposes or for other compatible purposes, and the procedure followed is also otherwise such that data concerning a given individual are not revealed to outsiders"* (response to Q3).

In **Germany** "*the rights of the data subject provided for in Articles 15, 16, 18 and 21 GDPR are limited to the extent that these rights are likely to make the realization of the research or statistical purposes impossible or seriously impair the restriction is necessary for the fulfilment of research or statistical purposes. In addition, the right to information according to Art. 15 GDPR does not exist if the data are necessary for purposes of scientific research and the provision of information would require a disproportionate effort*" (response to Q5).

In other countries, such as **UK, Italy and Sweden** there are no national law exceptions to the rules of Article 15-21 GDPR.[31]

Conclusion: Where possible and legally required, privacy notices should be drafted and made available to the data subjects, unless specific exemption applies. This includes preparing privacy information about processing of data in the context of BIGPICTURE which would be available on the platform.

Conclusion (2): Each Data Submitter must check whether they must enable data subjects to exercise their DSR rights. The consortium must agree on how the process of handling DSR will be carried out. In particular, to what extent the BIGPICTURE infrastructure can technically and legally enable the exercise of the DSR rights (e.g., deletion of the data).

1.2.3. Purpose limitation

The GDPR allows special considerations for situations when personal data to be used in a research project were originally collected for a purpose other than scientific research. This is the situation applicable to WSIs and Metadata in BIGPICTURE project. The Slides from which the WSIs are generated were collected for other purposes i.e. for most part are derived from clinical studies conducted by the Data Submitters or from provision of health services by them.

The basis for these considerations is related to the fundamental data protection principle of "purpose limitation" enshrined in Article 5(1)(b) GDPR. According to this principle personal data must be collected for specified, explicit and legitimate purposes and may not be further processed in a way incompatible with those purposes. There is an important caveat to this rule: further processing for, inter alia, scientific research purposes shall, in accordance with Article 89(1) GDPR, not be considered to be incompatible with the initial purpose. Further explanation of purpose limitation principle is provided in recital 50 of the GDPR. The full wording of those provisions is cited below.

This is often referred to as "presumption of compatibility" under which "*in principle personal data collected in the commercial or healthcare context, for example, may be further used for scientific research purposes, by the original or a new controller, if appropriate safeguards are in place"*.[32]

[31] Exceptions from Article 14.5 may still apply.
[32] European Data Protection Supervisor Preliminary opinion of 6 January 2020 on data protection and scientific research, https://edps.europa.eu/sites/edp/files/publication/20-01-06_opinion_research_en.pdf

Article 5.1. (a) (b)
1. Personal data shall be: (a)processed lawfully, fairly and in a transparent manner in relation to the data subject ('lawfulness, fairness and transparency'); (b)collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes; further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes ('purpose limitation')
Article 89
Safeguards and derogations relating to processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes 1. Processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, shall be subject to appropriate safeguards, in accordance with this Regulation, for the rights and freedoms of the data subject. Those safeguards shall ensure that technical and organisational measures are in place in particular in order to ensure respect for the principle of data minimisation. Those measures may include pseudonymisation provided that those purposes can be fulfilled in that manner. Where those purposes can be fulfilled by further processing which does not permit or no longer permits the identification of data subjects, those purposes shall be fulfilled in that manner. 2. Where personal data are processed for scientific or historical research purposes or statistical purposes, Union or Member State law may provide for derogations from the rights referred to in Articles 15, 16, 18 and 21 subject to the conditions and safeguards referred to in paragraph 1 of this Article in so far as such rights are likely to render impossible or seriously impair the achievement of the specific purposes, and such derogations are necessary for the fulfilment of those purposes. 3. Where personal data are processed for archiving purposes in the public interest, Union or Member State law may provide for derogations from the rights referred to in Articles 15, 16, 18, 19, 20 and 21 subject to the conditions and safeguards referred to in paragraph 1 of this Article in so far as such rights are likely to render impossible or seriously impair the achievement of the specific purposes, and such derogations are necessary for the fulfilment of those purposes. 4. Where processing referred to in paragraphs 2 and 3 serves at the same time another purpose, the derogations shall apply only to processing for the purposes referred to in those paragraphs.
Recital 50
(50)The processing of personal data for purposes other than those for which the personal data were initially collected should be allowed only where the processing is compatible with the purposes for which the personal data were initially collected. In such a case, no legal basis separate from that which allowed the collection of the personal data is required. If the processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller, Union or Member State law may determine and specify the tasks and purposes for which the further processing should be regarded as compatible and lawful. Further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes should be considered to be compatible lawful processing operations. The legal basis provided by Union or Member State law for the processing of personal data may also provide a legal basis for further processing. In order to ascertain whether a purpose of further processing is compatible with the purpose for which the personal data are initially collected, the controller, after having met all the

> requirements for the lawfulness of the original processing, should take into account, inter alia: any link between those purposes and the purposes of the intended further processing; the context in which the personal data have been collected, in particular the reasonable expectations of data subjects based on their relationship with the controller as to their further use; the nature of the personal data; the consequences of the intended further processing for data subjects; and the existence of appropriate safeguards in both the original and intended further processing operations.

On the basis of these provisions, and supported by the local law, some countries allow that - under defined conditions - the medical data collected for other purpose may be re-used for research without the explicit consent of the patients.

According to the country reports, this may be possible in **Belgium**, **Germany, Italy, UK, Sweden** and **Finland.** In **Netherlands** however patient "opt-in" (affirmative consent) is usually required. Only in limited cases the consent provided earlier may be sufficient to justify a new research project requiring additional data use. However, "*this needs to be reviewed on a case-by-case basis. Usually, some level of review of the anticipated research proposal will be needed (e.g., by a data access committee). This review will also verify whether the intended reuse of data matches the original purpose of the data collection (including wording of the consent which was provided earlier)"* (response to Q3).

The conditions that allow data re-use vary greatly between the Data Submitter's jurisdictions. Usually, they are connected with obtaining the approval of an ethics committee (as further explained in section 3 Ethical requirements below) and providing a guarantee of appropriate safeguards. Examples are provided below, however country reports should be consulted for further information.

In particular in the **UK** "*Data can be reused for scientific research provided consent or exemption (based on the Control of Patient Information Regulation and Section 251 of the NHS Act 2006, for e.g., in case of pandemic, for public concern, in regard to specific cancer types, for secret research) is in place and a Research Ethics Committee review is favourable"* (response to Q4).

Safeguards may be specified in detail or be left to the discretion of the researcher. In particular for **Finland** the law provides for a non-exclusive list of examples of the safeguards, such as:

- *"measures that enable subsequent checking and verification of the identity of the person who has recorded, altered or transferred personal data;*
- *measures to improve the competence of the personnel processing personal data;*
- *designation of a data protection officer;*
- *internal measures by the controller and the processor for preventing access to personal data;*
- *pseudonymisation of personal data;*
- *encryption of personal data;*
- *measures that ensure the ongoing confidentiality, integrity, availability and resilience of processing systems and services related to the processing of personal data, including the ability to restore the availability of and access to personal data in a timely manner in the event of a physical or technical incident;*
- *a process for regularly testing, assessing and evaluating the effectiveness of technical and organisational measures for ensuring the security of the processing;*
- *specific rules of procedure for ensuring compliance with the Data Protection Regulation and this Act when personal data are transferred or processed for another purpose;*
- *a data protection impact assessment in accordance with Article 35 of the Data Protection Regulation;*
- *other technical, procedural and organisational measures."* (response to Q5).

Also in **Belgium**, the list of general safeguards is provided in Title 4 Belgian Data Protection Law, in particular in Articles 190 to 204 thereof. They would only apply to a controller which "*may achieve his purposes only by derogating from the principles and obligations and rights of the data subjects*" and if no code of conduct, approved in accordance with Article 40 of the GDPR is respected.

Conclusion: Selection of an appropriate legal basis for the submission of the Data must be tackled on a case-by-case basis for each Data Submitter separately, taking into account compliance with the national law and practice on re-use of health data for research purposes and that these WSIs may be made available to other researchers in line with the project's objectives.

1.2.4. Data minimisation

Under the rule of data minimalization data collected and used in the BIGPICTURE project must be as minimal as possible, taking into account the intended purposes. This implies that:

- Data should not be held for further use, unless this is essential for reasons that were stated in advance.
- Data should only include as much data as required to successfully answer the research question(s).
- Data collected for one purpose cannot be repurposed without further consent. However, certain exceptions to this rule apply, as discussed under section 1.2.3 Purpose limitation above.

It has been observed that there is certain tension between complying with the principle of data minimalization and "reaping the benefits of <<big data>>."[33]

In the context of BIGPICTURE, the discussions about scope of the data needed evolve around Metadata which is to be accompanying the images. The current outcome of those discussions is provided in *D3.2 Minimal dataset deliverable.* Notwithstanding the requirements of the minimal Metadata, Data Submitters can at their own discretion provide additional Metadata, on top of what is defined as "minimum required" by BIGPICTURE. In such cases, they would need to similarly motivate the necessity for the increased scope, outside of the needs given by BIGPICTURE.

Conclusion: Discussions and agreement regarding the minimal dataset accompanying the WSI must take into account the rule of data minimization, in particular be aligned with research objectives for the creation of the repository. Reasons for the scope of the data should be documented.

1.2.5. Accuracy

According to the accuracy principle, personal data must be accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay.

The principle of accuracy is closely linked to the quality of data. It is argued that it applies to data itself, and not to further decisions made using the data (inferences drawn from that data).[34] This implies that BIGPICTURE must take and document steps to ensure that WSIs are of representative quality (as agreed by the relevant partners) and contain accurate Metadata. There should also be a process of removing and correcting incorrect or misleading data. This will also play a role when a data subject would exercise their DSR to rectify personal data. In order to keep the data accurate and up-to-date, data subjects should have the opportunity to update their personal data when it is inaccurate. However, some jurisdictions may allow exceptions for DSR - see section 1.2.2 above.

[33] Finck, Michèle & Biega, Asia. (2021). Reviving Purpose Limitation and Data Minimisation in Personalisation, Profiling and Decision-Making Systems, https://arxiv.org/ftp/arxiv/papers/2101/2101.06203.pdf

[34] Study: How the GDPR changes the rules of scientific research, https://www.europarl.europa.eu/RegData/etudes/STUD/2019/634447/EPRS_STU(2019)634447_EN.pdf

BIGPICTURE is taking significant effort in ensuring quality of its data. In particular the work will include developing standards and guidelines for generating WSI data and establishing a European Digital Pathology Quality Control Centre (T5.7).

Conclusion: BIGPICTURE must take and document steps to ensure that WSIs are accurate, of appropriate quality and contain accurate Metadata. There should also be a process of removing and correcting incorrect or misleading data, in particular at the legitimate request of a data subject (unless an exception applies).

1.2.6. Storage limitation

Under the storage limitation principle, personal data must 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. However, 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) GDPR subject to implementation of the appropriate technical and organisational measures required by the GDPR in order to safeguard the rights and freedoms of the data subject. This does not however imply that data stored for scientific research may be kept forever. Data Submitters (and accordingly data subjects) must be informed about the retention period, or at least its basis and rationale. After this period lapses, the data must be removed or anonymized.

Conclusion: Data Submitters and other controllers within the BIGPICTURE project must define a retention period for the pseudonymized WSIs and their associated Metadata that are collected for the participation in BIGPICTURE and/or available in the repository.

1.2.7. Integrity and confidentiality

The obligation described in this point means that personal data must be protected by appropriate technical and organisational measures to ensure its confidentiality, integrity and availability, as provided in Article 32 GDPR. This is closely related to requirements to provide appropriate safeguards under Article 89(2) GDPR. Moreover, the country experts were asked to identify data localization and infrastructure requirements (Q18-20).

1.2.7.1. Security requirements

Article 32 GDPR concerns the security of the processing. The controller must implement appropriate technical and organizational measures to ensure an appropriate level of security according to the risks posed by the processing. For this, the state of the art must be considered, as well as the costs of implementation and the nature, scope, context and purposes of processing, including the risk of varying likelihood and severity for the rights and freedoms of natural persons. Measures that can be taken include:

- the pseudonymization and encryption of personal data;
- the ability to ensure the ongoing confidentiality, integrity, availability and resilience of processing systems and services;
- the ability to restore the availability and access to personal data in a timely manner in the event of a physical or technical incident;
- a process for regularly testing, assessing and evaluating the effectiveness of technical and organizational measures for ensuring the security of the processing.

If the controller hires a processor, such processor must provide guarantees to implement security measures (Article 28(2) GDPR). The processor is responsible to take all measures pursuant to Article 32 and may be held liable for infringement of this obligation. Whether a security level is appropriate depends on the precise risks presented by a given processing operation. These can range from

accidental or unlawful destruction, loss and alteration, to unauthorized disclosure of, or access to, personal data transmitted, stored or otherwise processed.

Controllers and processors must ensure their employees handling personal data do not process them except as instructed.

This general obligation for ensuring the security of personal data will apply to processing of WSIs and associated Metadata in the context of BIGPICUTRE, including to storing them and making them available through the dedicated infrastructure. Most country experts (the **Netherlands, Sweden**) have indicated that there are no other specific rules applicable to security features of infrastructure for storing WSIs, other than the GDPR rules described above. However, based on local practice for the **Netherlands**, *"Usually, an encrypted connection is used for data transfers to meet to these requirements (…) For personal data these are usually interpreted in such a manner that some form of well accepted audit certificate (ISO 27001 or similar national certifications NE7510, NEN7512, NEN7513) is needed to prove that the security requirements are in place. This requirement for certification is usually also included in data processing agreements related to health data"* (responses to Q18 and 20*).*

In **Finland** "*From 1.5.2022 onwards, unless permanently anonymised, data can be made available only to processing environments verified and approved in accordance with the Secondary Use Act. Permissions granted now will only apply until 30.4.2022. Human biological samples in registered biobanks, technical records of the samples (like WSI) and associated data (sample related metadata, but also related diagnoses etc.) are available on application for research projects from the biobanks. The Biobank Act is expected to be renewed, and it is possible that the renewal will include similar restrictions on processing environments as the Secondary Use Act (see above)"* (response to Q2*).*

Conclusion: **BIGPICTURE must be mindful of the security requirements which are provided in the GDPR and local legislation, to ensure that Data Submitters from those countries can participate. The description of those safeguards should be made available to Data Submitters. In turn, each Data Submitter must check whether the safeguards offered by BIGPICTURE fulfil their requirements.**

1.2.7.2. Appointment of a DPO

According to Article 37 GDPR, a data protection officer ("**DPO**") must be designated if - among other circumstances - if the core activities of the controller or the processor consist of processing on a large scale of special categories of data and personal data relating to criminal convictions and offences. Within a group of entities, a single DPO can be assigned. Even when not required to do so, a DPO may be designated.

Some country experts have indicated national provisions requiring appointment of a DPO. For example, in **Belgium** *"There is also an obligation in the Belgian Data Protection Act to designate a DPO, specifically in cases where personal data are processed for scientific research purposes and the processing may result in a high risk. When personal data is processed for scientific purposes, the controller must anonymize or pseudonymize the personal data after it is collected. In cases of further processing, it is possible to de-pseudonymize the personal data only when necessary for the research purposes and, where applicable, after consulting the DPO. The DPO also must issue opinions on the use of the various pseudonymization and anonymization methods employed. The DPO thus provides an additional layer of security for the storage and processing of health data"* (response to Q20).

The DPO must demonstrate certain professional qualities and expert knowledge of data protection law. It may be a staff member or someone working through an external service contract. The contact details of the DPO must be communicated to the supervisory authority.

According to Article 38 GDPR, the DPO must be involved, properly and in a timely manner, in all issues which relate to the protection of personal data. The DPO is supported in its tasks by the controller and processor and must be able to work without instructions. Data subjects may contact the DPO with issues, as well as for the exercise of their rights.

The DPO must, according to Article 39 GDPR, inform and advise the controller or the processor; monitor compliance; provide advice where requested; cooperate with the supervisory authority; and act as the contact point for the supervisory authority. The DPO must have due regard to the risk associated with processing operations, considering the nature, scope, context and purposes of processing.

In the context of BIGPICTURE, the role of data protection officer is taken by BIGPICTURE GDPR taskforce. However, all organizations hosting the repository must confirm that they have appointed a DPO and contract details of these DPOs are made available to all controllers and processors involved in the project.

Conclusion: **The Data Submitters should be required to confirm that they appointed a DPO and if not - that they are exempted from this requirement. All other partners hosting the repository should be required to do the same. Details of DPOs should be exchanged and available for the controllers and processors of BIGPICTURE data.**

1.2.8. Accountability

Data controllers within BIGPICTURE must be able to demonstrate compliance with the GDPR principles. While it is not possible to prepare an exhaustive list of documents which are required from BIGPICTURE to achieve this goal, for instance, the following steps should be taken.

1.2.8.1. Record of processing activities

This will entail keeping a record of processing activities, which should include:

- the categories of data subjects from which the data are collected;
- types of data which are collected; what other parties the data is shared with, and, if applicable, details of any transfers of personal data to a third country (i.e., outside the EU);
- the time limits for erasure;
- and a general description of security measures.

1.2.8.2. Roles of the involved parties

Moreover, under the general umbrella of accountability responsibility for addressing obligations under the GDPR with respect to handling WSI should be assigned to involved parties.

The personal data processing operation must be conducted under the responsibility of a **controller**. A controller is, according to Article 4(7) GDPR "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." If two or more entities participate in the determination of the purposes and means of a processing operation, they would qualify as **joint controllers**.

A controller may use a **processor**, defined by Article 4(8) GDPR as "a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller." Given the importance of these two actors, the European Data Protection Board (EDPB) provided further guidance on this matter in Guidelines 07/2020 on the concepts of controller and processor in the GDPR ("**Guidelines 07/2020**").[35]

First, a data controller is usually the organisation as such, and not an individual within the organisation (such as the CEO, an employee or a member of the board), that acts as a controller.

Second, the controller decides on the key elements i.e. determines the purposes (why?) and means (how?) of the processing operation. The controller must decide on both purposes and means. However, some more practical aspects of implementation ("non-essential means") can be left to the

[35] Guidelines 07/2020 on the concepts of controller and processor in the GDPR, https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-072020-concepts-controller-and-processor-gdpr_en

processor. Importantly, the controller does not necessarily need to access the data to be qualified as a controller.

Third, as processing operations are becoming exceedingly complex, it is possible that multiple entities can be designated as joint controllers. As noted by the EDPB guidelines, joint participation can take the form of a *common decision* taken by two or more entities or result from *converging decisions* by two or more entities, where the decisions complement each other and are necessary for the processing to take place in such a manner that they have a tangible impact on the determination of the purposes and means of the processing. An important criterion is that the processing would not be possible without both parties' participation in the sense that the processing by each party is inseparable, i.e. inextricably linked.[36] Joint controllers should determine and agree on their respective responsibilities for compliance with the obligations under the GDPR in a binding document.[37]

It is, however, important to distinguish the case where controllers act jointly for the same purpose from the situation where multiple entities each pursue their own means and purposes. As noted by the EDPB, the fact that several actors are involved in the same processing does not mean that they are necessarily acting as joint controllers of such processing.[38] For example, exchange of a set of data without jointly determined purposes or means of processing should be considered a transmission of data between two independent controllers. Using a shared infrastructure by different controllers does not necessarily entail joint controllership. On the other hand, existence of joint responsibility does not necessarily imply equal responsibility of the various operators involved in the processing of personal data. The Court of Justice of the European Union (CJEU) in recent rulings[39] clarified that those operators may be involved at different stages of that processing and to different degrees so that the level of responsibility of each of them must be assessed with regard to all the relevant circumstances of the particular case.

Forth, regarding the **processor**, the GDPR assumes this entity to act under delegation by the controller and cannot process the data otherwise than according to the controller's instructions. Such delegation should be governed by a controller-processor agreement (data processing agreement, DPA), construed in compliance with Article 28 GDPR. As noted, while the processor may have some leeway to determine the non-essential aspects of the means of the processing, they should have no say over the purposes of the processing.

In the context of research projects, the Guidelines 07/2020 state that *"Several research institutes decide to participate in a specific joint research project and to use to that end the existing platform of one of the institutes involved in the project. Each institute feeds personal data it already holds into the platform for the purpose of the joint research and uses the data provided by others through the platform for carrying out the research. In this case, all institutes qualify as joint controllers for the personal data processing that is done by storing and disclosing information from this platform since they have decided together the purpose of the processing and the means to be used (the existing platform). Each of the institutes however is a separate controller for any other processing that may be carried out outside the platform for their respective purposes."*

Given the above, in order to impose the proper structure of data protection framework in the BIGPICTURE project is it crucial to carry out an assessment of roles of all consortium partners. This assessment must be made on a factual, rather than a formal, analysis of their actual influence with respect to determining the purposes and means of the collecting the WSIs, their storage and use for the purposes of the project. This does not however imply that all the partners of the BIGPICTURE consortium must assume such a role. Certain partners may not make any decisions regarding WSIs nor process such data and thus would not have a role defined under the GDPR.

Conclusion: It is crucial for BIGPICTURE to map the responsibilities of all involved parties and allocate controller, joint controller and processor roles based on those responsibilities.

[36] Guidelines 07/2020, pg. 3.
[37] Article 26 GDPR.
[38] Guidelines 07/2020, pg. 24.
[39] Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein v Wirtschaftsakademie, (C- 210/16), Tietosuojavaltuutettu v Jehovan todistajat — uskonnollinen yhdyskunta (C-25/17), Fashion ID GmbH & Co. KG v Verbraucherzentrale NRW eV (C-40/17).

Once this is determined, the policy framework and data sharing agreements should follow to ensure clear allocation of the obligations and rights with respect to the collection, storage and sharing of WSIs.

1.2.8.3. DPIA requirements

Article 35 GDPR requires a prior assessment of the impact of the envisaged processing operations on the protection of personal data, particularly when using new technologies and when the processing – taking into account its nature, scope, context and purposes – is likely to result in a high risk to the rights and freedoms of natural persons. The data protection officer may be requested advice.

According to the GDPR, a data protection impact assessment ("**DPIA**") is required when the processing concerns an automated processing providing systematic and extensive evaluation of personal aspects relating to natural persons and on which decisions are based that produce legal effects concerning the natural person; when processing on a large scale special categories of data; or when systematically monitoring a publicly accessible area on a large scale. A public list of these operations is maintained by the supervisory authority.

A DPIA must contain:

- a systematic description of the envisaged processing operations and the purposes of the processing;
- an assessment of the necessity and proportionality of the processing operations in relation to the purposes;
- an assessment of the risks to the rights and freedoms of data subjects; and
- the measures envisaged to address the risks, including safeguards, security measures and mechanisms to ensure the protection of personal data and to demonstrate compliance, taking into account the rights and legitimate interests of data subjects and other persons concerned.

A DPIA is not required if the processing is prescribed by law and if a general impact assessment was already conducted in adopting that legal basis.

If the processing can result in a high risk, the controller should contact the supervisory authority. If the authority finds that the processing would infringe upon the GDPR, or that the risks are insufficiently mitigated, it may provide a negative advice to the processing. Member States may allow authorities to require controllers to obtain prior authorization.

In the context of BIGPICTURE, the project combines health data from may sources and thus processing on a large scale special categories of data is likely to take place. Moreover several country experts have underlined the importance of conducting a DPIA. In particular in **Finland** "*If the data are considered personal data, then under the Data Protection Act (DPA), if GDPR's data subject rights (Art 15, 16, 18 or 21) are deviated from, a GDPR data protection impact assessment (DPIA) must be performed by the data controller of the research data and provided to data protection authorities in advance or relevant GDPR-approved code of conduct followed (see also answer 5 below). In addition, in the general BigPicture context, biobanks who are expected to provide the WSI should have performed DPIAs for their operations, but if they would start using BigPicture entities as processors to hold data for onwards disclosures they should perform a specific DPIA regarding this process (in accordance with the EDPB guidance and the conditions of Art. 35 GDPR)"* (response to Q3).

Also in **Italy**, conducting and making a DPIA publicly available is one of the conditions for taking advantage of an exception from seeking the consent of the data subject for processing of health data for scientific research purposes (see response to Q3). In **Belgium**, if the DPIA was appropriate and "*the controller processes sensitive data, within the meaning of article 9.1 of the Regulation, for scientific or historical research or statistical purposes*", this assessment should be added to the record of processing activities as one of the general safeguards (response to Q6).

BIGPICTURE will be conducting several (4) iterations of a Data Privacy Impact Assessment (D5.13-16).

Conclusion: BIGPICTURE should conduct a DPIA exercise to evaluate all relevant aspects of sharing of WSIs within the project, in line with the GDPR requirements. This DPIA should become a reference source that will feed into and contribute to own assessments conducted by the Data Submitters (acting as controllers).

1.3. Specific categories of health data

The legal experts have been asked to provide specific comments regarding the following specific types of WSIs:

- Containing genetics data
- Relating to rare diseases and Covid19
- WSI which were obtained in the context of clinical trials.

The responses which were obtained can be summarized as follows:

1.3.1. Data containing genetics data

Most of the experts did not indicate any particular restrictions relating to genetics data i.e. indicated that processing of genetic data in the same way as other health data. This is the case in **Belgium, Finland** (where however Genome Act is under preparation)**, the Netherlands (**although unsolicited findings resulting from genetics research would need to be addressed under the planned Medical Tissue Act, Wzl).

In **Germany**, the specific law *Gendiagnostikgesetz* includes several restrictions, *"but it explicitly states that these restrictions are not applicable to research use of genetic data. If this use is included in the informed consent, there are no restrictions for research use. If there is no informed consent, it is expected that genetic information, in particular germline information, will be restricted by the ethical committee"* (response to Q11).

Moreover, in **Italy**, there are specific rules which regarding processing and *"custody and safety*" of genetic data and biological samples. Special laws or provisions on genetic data also apply in **Sweden** and the **UK** (Human Tissue Act).

1.3.2. Data relating to rare diseases and COVID-19

Most of the experts (**Belgium, Finland, Germany, the Netherlands, Sweden, UK)** did not indicate any particular legal restrictions. However, it was emphasized that rare diseases data can be more identifiable despite protections (making it difficult or impossible to de-identify the WSIs) and this may affect decisions to grant access to them. For **Italy**, the expert identified certain guidelines regarding use of data on COVID-19, which should be consulted.

1.3.3. Data which were obtained in the context of clinical trials

For clinical trials, the response which was most common (e.g., **Finland, the Netherlands, UK)**, was that participating in a clinical trial requires a study-specific informed consent. Thus, it would need to be checked whether the original consent encompasses use of data for other research studies, such as BIGPICTURE.

Moreover, for **Belgium**, it was noted that "*if the human body material is used for purposes other than those described in the clinical trial dossier, then the material must be transferred to a biobank*", where it falls within the scope of the Human Material Body Law (response Q15).

Some countries (e.g., **Italy**) did not indicate any additional considerations.

1.3.4. Data categories under special considerations

In the **UK** there are specific data categories which may need to be excluded from the Data. This is

likely to apply in very narrow circumstances, given the objectives of the project. In particular: "*Data about abortions, certain sexual health conditions (including but not limited to HIV and Female Genital Mutilation) are held separately from the record by statute and would require special review from RECs and approvals from specialist centres (specialist treatment centres for conditions like HIV or interventions relating to FGM) to be shared"* (response to Q25).

Moreover, in the **Netherlands** "*usage of the national social security number (BSN) is strictly regulated. Health providers are required by law to use it in clinical care records, while its usage in health research is strictly forbidden"* (response to Q22).

Conclusion: BIGPICTURE should carefully consider whether to mandate the Data Submitters to provide WSIs with accompanying Metadata on specific categories of health data (genetic data, rare diseases, COVID-19). Provision of this Metadata should be marked as optional, which may be provided that the Data Submitter may share this information in line with the national law. WSIs containing such data in their Metadata may need an additional layer of ethical approval or special consideration. For WSIs derived from clinical trials, the Data Submitters must assure that the submission of the WSI is in line with any consent obtained from the individual and/or information provided to them during clinical trial.

Conclusion (2): Data Submitters should confirm that the Data that they make available to BIGPICTURE do not contravene any of their national laws restricting the disclosure of specific types of personal or health information, such as national ID number or information about certain diseases.

2. Non-clinical WSIs

The inquiry sent out to the legal experts included questions regarding rules which regulate use of WSI of animal samples (non-clinical WSI). In their responses, most of the experts indicated legislation on the use of laboratory animals and ethical requirements for animal testing/experiments. However, **no specific rules** were identified that regulate the creation of digital representations of animal Slides and use of WSIs of animal samples in such context as BIGPICTURE. For the **UK** the respondent advised that "*Should be checked with the specific site. There may be restrictions based on use of specified purpose, including standard research ethics guidelines"* (response to Q10).

3. Ethical requirements

The rules regarding ethical approvals vary greatly between the Data Submitters' jurisdictions. In many countries those rules are not codified but rather depend on local practice and the type of Data Submitter. For instance, for the **UK:** "*There is divergence in the rules for obtaining the approval depending on the authority. There are more than 80 different RECs, which vary in their expertise"* (response to Q2). Also in the **Netherlands**, "*As for WSIs, there is not a specific regime. In practice, biobanks and cohorts typically need to seek approval from an ethical board before initiating such a collection to meet the requirements in the Medical Treatment Contact Act (WGBO). For other data collections several institutions have so-called Institutional Review Boards (IRB), which is not a legal obligation but an institutional safeguard that all procedures are followed. There is clearly a lack of consensus in this area"* (response to Q9).

If applicable, typically ethical requirements entail a specific regime of obtaining an authorization from the ethics committees to access health data and use it for research purposes.

It is important to note that **all of the countries in question** indicated that the Data Submitter would be required to obtain an ethical approval to contribute Data to the project. In particular, in **Germany** *"For BIGPICTURE, these existing approvals would need an amendment to include WSIs as well as transfer to the repository. For retrospective projects of one single department and use of already existing data (including slides and WSIs) with anonymisation, a simplified ethical board approval is required"* (response to Q9). Similarly, this is the case in **Finland** "*For access to biobank collections, an ethical review is also required, but this is sometimes performed by the biobanks' own scientific*

boards as part of processing the access application. However, in practice in many (if not most) cases a statement from a statutory committee established in accordance with the Medical Research Act is required (requires a separate prior application)" (response to Q9).

Furthermore, in the context of BIGPICTURE, while the organization of the honest broker mechanism is still being discussed in the framework of WP3, the ethical requirements for obtaining the ethical approval should be further investigated and aligned with the planned establishment of data access committees ("**DAC**") . In particular, it must be discussed with the participating Data Submitters if under their authorizations they could agree to delegate to outside DACs the right to make decision on accessibility of contributed WSIs to other researchers. As the **Finnish** respondent noted *"Access to samples and data in the discussed context is legally controlled and permissions are needed. Providing these permissions cannot be freely delegated, so only the entities (biobanks and Findata) recognised in the law can grant access. BIGPICTURE partners or a legal vehicle created by the project would not be such an entity, unless it establishes itself as a biobank registered in Finland. This would most likely be possible, but there are uncertainties related to the pending changes and the new biobank act may affect the feasibility of this option"* (response to Q21).

Based on the received country responses it is apparent that there is no unified framework within the Data Submitters jurisdictions on contributing data to research repositories and also the rules of ethics; approval conditions are often vague. This is to be further discussed together with the actual initial Data Submitters organizations and the conditions upon they receive their ethical approvals for the participation in the project. From a high-level perspective, given the findings of the expert reports, the following scenarios should be further examined:

a) Data Submitter's ethical approval encompasses transfer of data to the repository and making the data further available to other researchers upon fulfilment of certain criteria (presumably to be examined by DAC) would not require a new ethics approval.

For example, as noted by report for the **Netherlands** "*In the context of BigPicture, since the project does not involve new research on human subjects, the external partners applying to the biobanks to access their submitted WSIs via the BigPicture infrastructure would not have to obtain an ethics approval"* (response to Q17).

b) Separate approval from the Data Submitter's ethical committee must be obtained by the researcher seeking the access to WSIs, before the access is granted.

This is illustrated by the following comment from the **Finland** expert "*BIGPICTURE doesn't seem like a medical research project as such, but rather an infrastructure to support many future activities, it doesn't seem possible to provide independent control to BIGPICTURE, but biobanks will make a decision on this if an application is made and a research plan presented (requires also an ethical review). In such case, generally the researcher's application made through this infrastructure to a contributing biobank should be treated as any other application to the biobank, so the researcher would need to fulfil the conditions for access to data. There could be however some room for discretion for the biobank for setting up "fast track" for such applications if they for example were able to be standardised to warrant expedited processing. Moreover, processing could be subcontracted to BIGPICTURE partners as long as the biobank controls access. This is under the assumption that data provider retains control over access to the WSI it contributes"* (response to Q32).

In **Belgium** the legal situation depends on various factors, including which law would apply for the researcher which seeks to obtain the WSI. Moreover, there is uncertainty on how the ethics committees would approach possible delegation of competence to allow access to WSIs to DACs. As noted by the expert team: "*The specific legal provisions applicable to the further use of the WSI slides, once made accessible via the BigPicture repository, will depend on 1) the data protection law applicable to the research project for which these slides will be used, and 2) if Belgian data protection law is applicable to this research project, the requirements of this particular research project with regard to the necessity to benefit from the research exemptions mentioned in Art. 89(2) of the GDPR (see the 4 possible scenarios in the answer to Q3). It is not excluded that ethical committees in Belgian academic hospitals will be reluctant to approve "blank" access to their WSI slide collections and to delegate their competence to an entity belonging to a common repository. It is also not excluded that*

ethical committees of academic hospitals in Belgium will be reluctant to approve the use of WSI slides in research projects regulated by data protection laws of other Member States." (response to Q32).

Conclusion: Each Data Submitter must confirm that they comply with the rules applicable to seeking and adhering to ethical approval in their jurisdiction.

Conclusion (2): BIGPICTURE must engage in discussions with concrete Data Submitters on the approach of their ethical committees to delegating authorization of access to WSIs to a DAC (if appointed outside of their organization).

4. Specific regulations regarding biobanks

As demonstrated by the expert responses, biobank law is not harmonized between the jurisdictions of the Data Submitters, located in the European Union and the UK. Furthermore, not all countries in scope have specific rules relating to biobanks.

In particular, specific biobank law is implemented in **Finland**, **Belgium** and **Sweden.**

No such law was identified in the **Netherlands, Germany, Italy** and the **UK**. However, there exist certain national rules on the use of human materials (tissues and blood). For example, in the **Netherlands** rules on use of human material are provided in the Medical Treatment Contract Act (WGBO) and Dutch Code for proper secondary use of human tissue. They contain "*concrete standards for Dutch health research that are accessible to researchers*" (response to Q16). In the **UK**, there is fragmented patchwork of laws that applies to medical research on humans, providing for a distinction between human material (samples) and data. UK's Human Tissue Act would not apply to WSIs, though (response to Q17).

The further conclusion which can be drawn from the received input is that for most of the jurisdictions, it is questionable whether the WSIs will be under the scope of biobank law. In particular, in **Belgium** "*Images of human body samples are "data" ("personal data" if the person concerned is identifiable). A collection of such images is not a biobank, according to the Q & A in the Compendium*" (response to Q3). Also in **Sweden** it is "*the common interpretation of the Biobank Act has been that storing tissue samples in biobanks does not amount to processing of personal data as such*" (response to Q16). Thus, **Swedish** Biobank Act would not be directly applicable to sharing the digital images of those samples (WSI). This is also the case in the **UK**, in the sense that distinction is made between samples and data.

The situation is different in **Finland**. Here, there is a different regime for accessing data under Secondary Use Act and rules which apply to biobanks, as they are regulated under the Biobank Act (response to Q6). However, in principle, the WSI would be considered as falling under the scope of Biobank Act if obtained from the biobank: *"Biobanks maintain collections of samples and related data (metadata, clinical data etc. linked to the samples) for future research purposes. WSI produced by the biobank would be considered a technical record of the sample*¨(response to Q16). Consequence of this is a requirement for the material/data transfer agreement (M/DTA) to be concluded to access a WSI from a biobank in Finland.

The biobank acts can have further consideration on the set up of the roles and responsibilities of BIGPICTURE partners. As noted by **Finnish** expert: "*It would seem that the only way Finnish data can be in BIGPICTURE is under a subcontract for processing, with no independent decision making over disclosures or transfer. Another option could be registering as a biobank in Finland."* Furthermore: "*If BIGPICTURE however wanted to have independent control of the Finnish WSIs (e.g., by granting access requests to the researchers), this would seem possible only if the controlling entity/entities registers a biobank in accordance with the Finnish Biobank Act and concludes agreements with the original provider biobanks. This should be further considered by the project partners"* (response to Q22).

Moreover, the biobank laws would apply to the transferring of the glass Slides as such. While this issue was outside of the scope of the assessment, nevertheless some reports contain additional information about transferring the original Slides (e.g., **Sweden, UK**) or more detailed analysis of

biobank law was provided (**Belgium**).

Conclusion: Each Data Submitter must indicate which special rules apply to the submitted Data under biobank laws. In particular if there is a requirement for concluding a material transfer agreement (MTA).

5. Intellectual Property (IP) rights to the Data

The possible IP rights to the WSI and accompanying Metadata include:

- database rights
- copyright and photograph protection,
- protection of trade secrets.

Below, we summarize the essence of each type of right and provide a summary of the findings from relevant jurisdictions.

5.1. Database rights

Databases are protected under EU Law. The Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases ("**Database Directive**") was adopted in 1996 and provides for uniform legal protection regime in the EU.

The Database Directive protects collections, sometimes called 'compilations', of works, data or other materials which are arranged, stored and accessed by means which include electronic, electromagnetic or electro-optical processes or analogous processes.[40] Protection extends to collections of independent works, data or other materials which are systematically or methodically arranged and can be individually accessed.[41]

In accordance with the Database Directive, the Member States were obligated to provide for a right for the maker of a database which made investments in creation of the database. The objective of this *sui generis* right (specific property right for databases) is to ensure protection of any investment in obtaining, verifying or presenting the contents of a database for the limited duration of the right.[42] This right encompasses the right to prohibit extraction and re-utilization of the database or its substantial part.[43] This right may be transferred or licensed and applies irrespective of the eligibility of that database for protection by copyright or by other rights and irrespective of eligibility of the contents of that database for protection by copyright or by other rights. The right of protection of databases is without prejudice to rights existing in respect of their contents. It runs for 15 years from the first of January of the year following the date of completion of the database. Moreover, the Database Directive protects databases by copyright if they are original by reason of the selection or arrangement of their content.

The rules of the Database Directive were transposed by the responding jurisdictions and may apply to the collection of the WSIs provided by the Data Submitters. For instance, in **Belgium** exists a separate regime for database rights in the Belgian Act on Economic Law. In **Germany**, such protection is provided in the *Urheberrechtsgesetz*. Also in **Italy**, the collection of WSI may be a database protected under one of the bases mentioned in the Italian Copyright Law. Detailed responses on the rules are provided in Q24 of the country reports.

However, as pointed out by some of the respondents, the consequences to collection of WSI in the context of BIGPICTURE would be rather limited, assuming that the Data Submitter was the holder of the database rights (e.g., maker of the database). For **Finland** it was noted that "*There are specific legal rules addressing how patient and biobank data can be and, in some cases, must be shared, so*

[40] Recital 13 Database Directive. Article 1.2 e provides that "database" shall mean a collection of independent works, data or other materials arranged in a systematic or methodical way and individually accessible by electronic or other means.
[41] Recital 17 Database Directive.
[42] Recital 40 Database Directive.
[43] Article 7(1) Database Directive.

the IPR doesn't seem particularly consequential to BIGPICTURE" (response to Q24). Also in **Italy**, *"If the hospital was the "maker" of the WSI database, it would be entitled to the "sui generis right" to this database and would be able to grant licenses to use these WSI in the research projects"* (response to Q24).

In the **Netherlands** it may be even doubtful if the rights to the database would arise *"Database rights could apply to an organized collection of slides, but technically, such database rights would apply to the infrastructure and the entirety of the database, not individual sides or individual contributions from the data providers"* (response to Q24).

5.2. Copyright protection and photograph rights

Copyright protection is only partially harmonized within the EU, with the regulatory framework for copyright and neighbouring rights (acquis) consisting of 11 directives and 2 regulations. There is however no universally recognized definition of copyrighted works.[44] It is generally accepted that it includes every original work of authorship, irrespective of its literary or artistic merit. The ideas in the work do not need to be original, but the form of expression must be an original creation by the author.[45]

In the context of BIGPICTURE, as indicated by some of the country respondents, a collection of WSIs may potentially be recognized as copyrightable work (**Italy**). Moreover, some of the jurisdictions (**Finland**, **Sweden, Belgium**) recognize *sui generis* right to photographs, even if the threshold of artistic work is not met. However, the latter right may be subject to additional conditions. For instance, in **Belgium** the right to photograph is included in Article XI.174 of the Belgian Act of Economic Law. However, "*the legal doctrine and jurisprudence unanimously declare that the person depicted must be recognisable. If the person depicted is not recognisable, this right to photograph is not applicable and the consent of the depicted person is no longer necessary"* (response to Q24). For this reason, the **Belgian** right to photograph would not be applicable for WSI, as the person will not be recognisable from the image.

In addition, should the IP rights to WSI and Metadata arise, they would most likely be vested in the Data Submitter (**Finland**), either by law or by employee contracts. For example in **Sweden:** *"The neighbouring rights to the photographs would typically be vested in the institution (employer) rather than individual employees, by way of employment contract"* (response to Q24).

Both the copyrights and database rights may be transferred or licensed to the project and the user. For the **UK** the respondent noted that *"Existing IP (specifically the IP relating to records) remains with the Health Care Trust / Health Board that are the Data Controllers of the Records. Arising IP is negotiated in a per case basis"* (response to Q23).

Other than this reservation, no further particular limitations in this respect were observed by the respondents. In particular for **Germany** *"The pathologist / hospital can provide the rights to use this image to the BIGPICTURE repository"* (response to Q24).

5.3. Protection of trade secrets

In the EU, following the Trade Secrets Directive[46] a trade secret means information which meets all of the following requirements:

- it is secret in the sense that it is not, as a body or in the precise configuration and assembly of its components, generally known among or readily accessible to persons within the circles that normally deal with the kind of information in question;
- it has commercial value because it is secret;

[44] Article 2 of Berne Convention for the Protection of Literary and Artistic Works (1886) states that: "The expression 'literary and artistic works' shall include every production in the literary, scientific and artistic domain, whatever may be the mode or form of its expression." The Convention lists examples of such works.
[45] WIPO Understanding Copyright and Related Rights (https://www.wipo.int/edocs/pubdocs/en/wipo_pub_909_2016.pdf)
[46] Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure (further as "**Trade Secrets Directive**").

- it has been subject to reasonable steps under the circumstances, by the person lawfully in control of the information, to keep it secret.[47]

The Trade Secrets Directive was transposed into laws of Member States and similar conditions have been provided for in those laws. The Trade Secrets Directive provides harmonization of the rules of civil law through which victims of trade secret misappropriation can seek protection. Trade secrets however are not intellectual property rights thus the holder of the trade secret does not enjoy an exclusive right. The Trade Secrets Directive does not alter the legal obligations requiring to divulge information for such public policy objectives nor regarding open access rights. Importantly, it does not provide any mandatory rules on what information cannot be disclosed by the rightful owner of such information.

As the scope of Trade Secrets Directive is quite broad, trade secrets protection may apply to WSI and Metadata. As the **Belgian** report states "*it can be concluded that the WSI (and the metadata connected to it) can be considered a trade secret. The Data Submitter could therefore decide to keep certain information secret and remove it from the WSI under trade secret protection*" (response to Q26). Also in **Finland** *"collections/compilations of WSI and related data could clearly be subject to database or catalogue protections, like any data/information, also personal"* (response to Q26).

Nonetheless, as noted by some respondents, the important condition of the trade secrets protection to apply is that the holder of the information must take reasonable measures to keep it secret. For instance, in **Belgium** the following measures may be considered "*surveillance in a vault, a password, encryption, etc.*" (response to Q26).

On the reverse side, this does not mean that the Data Submitter *must* keep any of its business information in confidence and thus must remove any of its commercially valuable information from the WSI or its Metadata. However, if this information is no longer kept secret then it would lose its protection. As put by the **Netherlands** expert *"There is no requirement to remove trade secrets in the Wbb [Dutch Trade secret act], this depends on the decision of the holder of the trade secret if they want to make it public"* (response to Q26).

Conclusion: The Data Submitters must check if IP rights to Data arise in their jurisdiction. If this is the case, Data Submitter must confirm that he obtained such rights, either by virtue of law or by contract (e.g., by an agreement with the researcher), and can submit Data in the scope required to participate in BIGPICTURE.

Conclusion (2): BIGPICTURE license/licenses should be drafted to contain license (or transfer of rights) for the use of the submitted Data for the purpose of the BIGPICTURE project. License requirements should also be included in the contract (or terms and conditions) of the use of Data by researchers accessing the repository.

6. Contractual limitations and open data (public data) restrictions

6.1. Contractual limitations

The majority of the respondents did not indicate any contractual limitations (other than requirements of the personal data protection, which were discussed above). For instance, the **Belgium** report states that there is no specific law in **Belgium** that regulates the use of WSI. However, the rules of use of health data, including requirement for pseudonymisation or anonymisation, apply (response to Q23).

Certain provisions however exist in **Finland.** In particular *"research results must be made public."* This is in addition to other requirements described in other parts of this report, such as keeping the data confidential and used only as necessary under the approved research plan (response to Q23).

In **Netherlands** the laws provide for narrower legal basis for use of data if it is used for research for commercial purposes. Namely *"Dutch implementing legislation for GDPR allows exceptions to research use without consent only for research performed in the public interest (art 24 and 28 UAVG).*

[47] Article 2(1) Trade Secrets Directive.

Consequently, research use for commercial purposes is only allowed based on consent" (response to Q23).

6.2. Open data rules

All of the questioned countries have identified rules on access to public sector information. For EU Member States they have been implemented by requirements of the directives on open data and the re-use of public sector information.[48]

None of the respondents have however indicated that those laws would limit the use of WSI in the context of BIGPICTURE. On the contrary, it was noted that the open data laws would not apply to WSIs. For example, based on responses to Q29-30 to the questionnaires:

- **Finland:** *"The lex generalis Act on the Openness of Government Activities applies to public sector operators including the biobank operators contemplated in BIGPITRURE, and to some extent also private biobanks. I do not foresee any relevant impact as access to the relevant data is specifically stipulated."*
- **Germany:** *"The rules of Access to Information (HDSIG) apply to universities and university hospitals insofar as they do not operate in the field of research and teaching."*
- **Italy:** "*The access to public information is regulated by the D.Lgs. n. 33, 14 March 2013, as modified by the D.Lgs. n. 97, 25 May 2016, (which implemented the Freedom of Information Act – FOIA -). However, this legislation does not add any restrictions or conditions apply to making the WSI available to the BigPicture."*
- **Netherlands:** *"As mentioned above, WOB rules are not applicable to BigPicture partners or slide contributors, as only public bodies are subject to them."*
- **Sweden:** *"The provisions on public access to information would however not apply to data submitter's own participation in a research project and its submission of data to such project. They would only be applicable if a third party researcher requested data from an institution, which was subject to public access of information laws."*

Conclusion: Based on the responses from the country experts, the existing public data laws should not be an obstacle in providing the Data to the project nor imposing conditions for their use in the context of BIGPICTURE. However, given the different types of Data Submitters which may enrol in the project, this should also be verified by the Data Submitter prior to WSI submission.

7. Other issues

Below we shortly summarize responses to remaining questions from the legal assessment questionnaire.

7.1. Digitalization of WSI

No specific rules regarding digitalization (beyond data protection laws) were identified and described. For instance in **Belgium** there is "*no specific law or specific provisions in the law of 19 December 2008 regarding digitalization of images of human tissue and samples in Belgium. Generating images from human tissue and digitalization of those images will be qualified as a processing of personal data if the patient is identifiable from whom the tissue or samples have been collected. Digital pathology slides will be considered as <<representations>> of HBM and, therefore, as <<data>> (<<personal data>> if the patient is identifiable)"* (response to Q8).

Some respondents also pointed to general rules on handling of human biological material (**Finland**). Those rules are for example in **Sweden (**response to Q8). In the **UK** rules of Human Tissue Act and

[48] The most recent Open Data Directive (Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information), entered into force on 16 July 2019, replacing the Public Sector Information (PSI) Directive.

UK GDPR provisions around Genetic Data must be followed.

7.2. Requirements for development of AI/ML

No requirements relating to artificial intelligence/machine learning in the context of WSIs were identified by the country respondents. Some respondents (for e.g., **Finland**) noted that EU's artificial intelligence regulation might be enacted in time to affect BIGPICTURE.

7.3. Honest broker mechanism

In some jurisdictions honest brokers services exist either provided for by the law, or in practice. In **Belgium** it is a legal requirement to use honest broker services in certain situations i.e. "*pseudonymisation by a trusted third party or by one of the data providers is mandatory before transmitting health data to researchers who wish to benefit from one or more of the exemptions mentioned in Art. 89(2) GDPR, if they collect health data from multiple sources"* (response to Q21).

In the **Netherlands** while there is no "official" honest broker service, "*there is a fairly commonly used commercial service for secure data linkage which has some features of an honest broker. This service is called ZorgTTP"* (response to Q21).

Moreover, in **Finnish** permissions authority under the Secondary Use Act, *Findata*, can be seen as type of honest broker, as it provides permissions and access to health and social care data from many data controllers and may "pre-process" data for the researcher, for example pseudonymize or anonymize it (response to Q21).

None of the other respondents identified any requirements of the honest broker mechanism that would impact the BIGPICTURE project.

7.4. General information about databases and registers containing patient's health data and biological samples

Each of the country experts provided overview of the registers which may contain data relevant for the project (responses to Q1). This information may be further analysed to obtain additional Data Submitters at later stages of the project.

7.5. Potential consequences of non-compliance

BIGPICTURE as a whole is not a legal entity, thus responsibility for any non-compliance with the above mentioned rules (e.g., infringement of the GDPR rules) would need to be evaluated towards particular partners or even - in some cases - to individuals within those organization (e.g., for criminal liability). The types of the potential sanctions vary between the involved countries (responses to Q7 to the questionnaires). Moreover, the GDPR contains very complex rules regarding the authority which could be authorized to investigate a privacy infringement. In practice, given different scenarios which may occur and also the transnational character of the consortium, the responsibility of BIGPICTURE members in case of any breach of the above mentioned rules would need to be ascertained on a case by case basis. It is not possible to draw an overarching principle at this stage, especially given that the roles of the partners are not defined. In particular, for data protection infringements, the scope of responsibility of each entity is related to the position it has regarding the data (controller, processor, joint controller). This is explained above in section 1.2.8.2.