DELIVERABLE 11.2

NEC – REQUIREMENT NO. 2

WORK PACKAGE NUMBER: WP11

WORK PACKAGE TITLE: ETHICS REQUIREMENTS

TYPE: REPORT

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)

Document Information

iReceptor Plus Project Information	
Action full title	Architecture and Tools for the Query of Antibody and T-cell Receptor Sequencing Data Repositories for Enabling Improved Personalized Medicine and Immunotherapy
Action acronym	iReceptor Plus
Grant agreement number	825821
Action coordinator	Prof. Gur Yaari
Action start date and duration	1st January 2019, 48 months
Action website	http://www.ireceptor-plus.com

Deliverable Information	
Work package number	WP11
Work package title	Ethics requirements
Deliverable number	D11.2
Deliverable title	NEC – Requirement No. 2
Description	Personal data are transferred from the EU to a non-EU country or international organisation, confirmation that such transfers are in accordance with Chapter V of the General Data Protection Regulation 2016/679, must be provided and submitted as a deliverable.
Lead beneficiary	time.lex
Lead Author(s)	Liesa Boghaert; Jos Dumortier
Contributor(s)	
Revision number	
Revision Date	

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)

Status (Final (F), Draft (D), Revised Draft (RV))	D
Dissemination level (Public (PU), Restricted to other program participants (PP), Restricted to a group specified by the consortium (RE), Confidential for consortium members only (CO))	CO (including Commission Services)

Document History

Revision	Date	Modification	Author
1	09.06.2019	Initial version	Jos Dumortier and Liesa Boghaert
1	24.06.2019	Review and edits	Gur Yaari
1	25.06.2019	Correct names	Milena Mirkis
1	26.06.2019	Comments	Artur Rocha
1	26.06.2019	Comments	Brian Corrie
2	28.06.2019	Final version	Jos Dumortier and Liesa Boghaert

Approvals

	Name	Organisation	Date	Signature (initials)
Coordinator	Prof. Gur Yaari	Bar Ilan University	30.06.2019	GY
WP Leaders				

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)

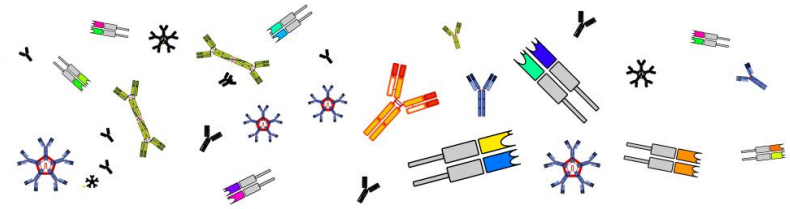
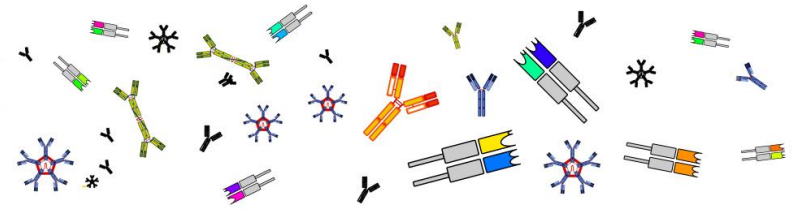


Table of Contents

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)



Executive Summary

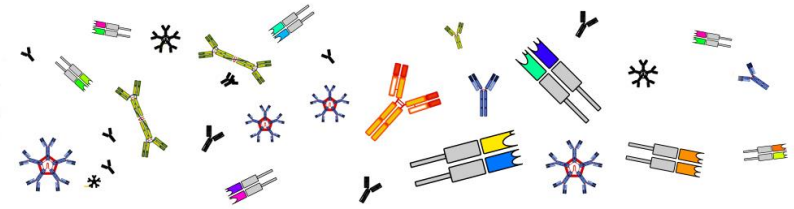
This deliverable is dedicated to confirming that any transfer of personal data from the EU to a non-EU country or international organisation is in accordance with Chapter V of the General Data Protection Regulation (GDPR).

To this aim, the deliverable first provides a short introduction to the objectives of the iReceptor Plus action and their relationship to the subject of the deliverable. The document then proceeds by setting out the elementary concepts of the GDPR and the different requirements the GDPR prescribes for a transfer of personal data to a third country or international organisation. In the following section of the deliverable, the data processing activities performed in the course of the action are analysed, so as to assess them in light of the requirements of the GDPR. From this assessment it appeared that the core data processing activities carried out in the context of this research and innovation action do not involve any processing of personal data at this stage, but rather makes use of anonymous data. This leads to the conclusion that the GDPR, including Chapter V of the GDPR, for now does not apply to core datasets exchanged in the iReceptor Plus research and innovation action.

Nevertheless, to mitigate the risk that throughout the research and innovation action personal data would nonetheless be processed at a later stage or that anonymous data could potentially be re-identified, the beneficiaries will (1) consider adhering to the BBMRI-ERIC GDPR Code of Conduct for health research and will (2) monitor and periodically re-assess the identifiability of the data that are processed in the course of the action, with the aim of ensuring strict compliance with the highest standards of data protection, as enshrined in the GDPR.

Such re-assessment will, in our opinion, in any case be necessary as of month twenty-four of the project.

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)



1 Introduction

1.1 Purpose and scope of the deliverable

As a part of Work Package 11 (Ethics requirements), Deliverable 11.2 aims at (1) analysing the processing activities performed in the iReceptor Plus action and (2) providing confirmation that any transfer of personal data from the EU to a non-EU country or international organisation throughout the action complies with the requirements set out in Chapter V of the General Data Protection Regulation[1] (GDPR).

In terms of scope, it is important to emphasize that this deliverable will only deal with data processing activities that are performed in the course of the action and will not consider any data processing that will be carried out once the iReceptor Plus research and innovation action is finalised and fully operational. This limitation of scope corresponds to the description of deliverables in the Grant Agreement.

1.2 iReceptor Plus and the GDPR

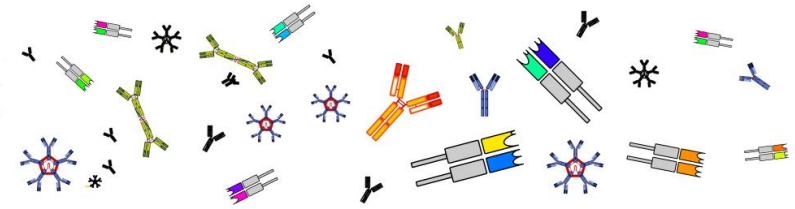
The objective of iReceptor Plus as a research and innovation action is to build a common scalable platform to integrate distributed repositories of Adaptive Immune Receptor Repertoire sequencing data (AIRR-seq data) for enabling improved personalized medicine and immunotherapy for diseases with an immune component. iReceptor Plus will be designed as a network of federated repositories that facilitates data queries and advances analyses through a centralized web portal (the iReceptor Plus Scientific Gateway). In essence, this means that iReceptor Plus will be a freely and openly available software platform, which enables (1) the querying of multiple repositories at once by common metadata searches, (2) the analysis of federated data aggregated from multiple repositories and (3) the integration of these data with other types of large scale human health and genomic data.

To develop this software, beneficiaries based both within and outside the European Union (EU) will bundle their expertise. Within the EU, beneficiaries from France, Germany, Spain, Portugal and Belgium are involved in the action. The non-EU beneficiaries in the action are based in Norway, Israel, Canada and the United States of America.

Given that the iReceptor Plus action will result in a software platform that facilitates access to and analysis of both AIRR-seq data and non-AIRR-seq data (such as clinical and biological data)

[1] 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).

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)



stored in distributed repositories, it is important to analyse what the exact processing activities during the development phase of the action are, and if these processing activities comply with the GDPR and more specifically, Chapter V of the GDPR on 'transfers of personal data to third countries or international organisations'.

The following sections of this deliverable will therefore:

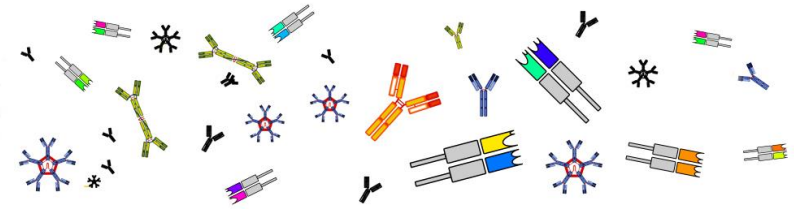
- **SET OUT** the requirements that the GDPR puts forward in relation to transfers of personal data to third countries[2] or international organisations (**section 2**);

- **ANALYSE** the data that will be processed by beneficiaries during the action as well as their transfer outside of the European Economic Area (EEA) (**section 3**);

- **ASSESS** the data processing activities in light of the requirements set out in (Chapter V of) the GDPR and **RECOMMEND** on any further steps to take (**section 4**).

[2] Third countries in the sense of the GDPR are countries other than the EU Member States and the three additional European Economic Area (EEA) countries Norway, Iceland and Liechtenstein that have adopted a national law implementing the GDPR.

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)



2 GDPR requirements for the transfer of personal data to third countries or international organisations

This section will start off with a short introduction to the GDPR, which will allow the reader to get a grip on some of the most elementary concepts determined by the GDPR and to have a better understanding of the following parts of this section, which deal with the requirements the GPDR sets forth when personal data are transferred outside of the EEA.

2.1 Introduction to the GDPR

The General Data Protection Regulation is an EU regulation[3] laying down rules relating to the protection of natural persons with regard to the processing of personal data and rules relating to the free movement of personal data.

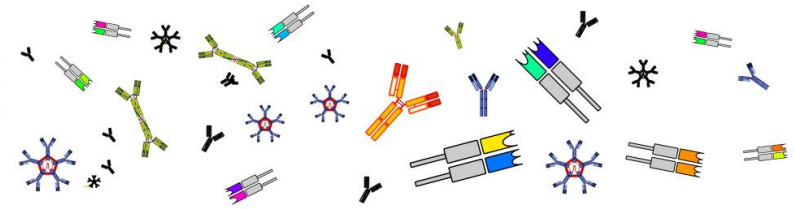
'Personal data' is defined by the GDPR as: *'any information relating to an identified or identifiable natural person **('data subject')*** (art. 4(1) GDPR).

An **'identifiable natural person'** or 'data subject' 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'.*

> It follows from these definitions that **anonymous data** are not covered by the GDPR. Indeed, the GDPR states that:
>
> '*The principles of data protection should therefore not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that a data subject is no longer identifiable. This Regulation does not therefore concern the processing of such anonymous information, including for statistical or research purposes.'* (Recital 26 GDPR)

[3] A regulation is a legal act of the European Union that is immediately and simultaneously enforceable as a law in all EU Member States. Regulations should be distinguished from directives, which need to be transposed into national law by the Member States.

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)



However, **pseudonymous data** remain subject to the provisions of the GDPR.

Pseudonymous data are personal data which have undergone pseudonymisation, but which can still be attributed to a natural person by the use of additional information. These data are considered to be information on an identifiable natural person in the sense of the 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'.* (art. 4(2) GDPR)

On a material level, the Regulation applies to the processing of personal data wholly or partly by automated means and to the processing other than by automated means of personal data which (are intended to) form part of a filing system (art. 2 GDPR).

Territorially, the GDPR applies when personal data are processed in the context of the activities of a controller or processor[4] established in the EU (regardless of whether the processing itself takes place in the Union or not), as well as to the processing of personal data of data subjects in the EU by a non-EU controller or processor when that processing relates to offering of goods or services, or the monitoring of behaviour that takes place in the EU (art. 3 GDPR).

Essentially, the GDPR:

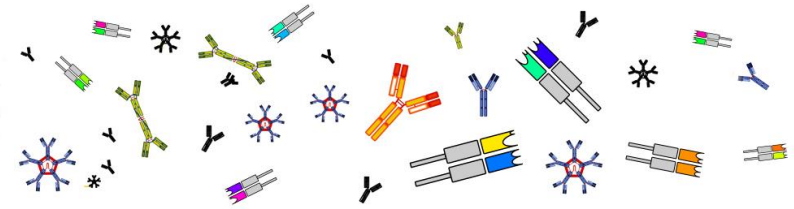
- Articulates the general principles that should be followed when personal data are processed (e.g. principle of data minimisation, purpose limitation, accuracy…),

- Puts forward the legal bases that may justify the processing of personal data (e.g. consent, legitimate interests, legal obligation…),

- Explains the rights that data subjects have in order to retain control over their personal data (e.g. right to be informed, right of access, right to erasure…),

[4] The controller is the natural or legal person, authority or body which alone or jointly with others determines the purposes and means of the processing of personal data. It is the one who decides *why* and *how* personal data are processed. The processor on the other hand, is the natural or legal person, authority or body which processes personal data *on behalf of* the controller.

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)



- Lays down the obligations of controllers and processors when processing personal data (e.g. technical and organisational security measures, record keeping…),

- Dictates the conditions under which personal data may be transferred to third (non-EU) countries or international organisations,

- Points out the remedies, liabilities and penalties that exist in case of breach of the GDPR, as well as the ways in which compliance will be supervised.

2.2 Chapter V of the GDPR: transfer of personal data to third countries or international organisations

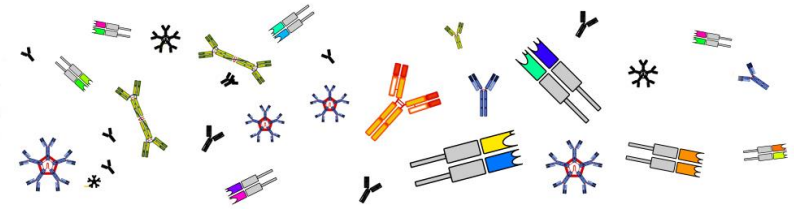
As mentioned above, the GDPR dictates the conditions under which personal data may be transferred to third countries or international organisations. Indeed, whenever personal data which are undergoing processing or are intended for processing are transferred to a third country or to an international organisation, Chapter V of the GDPR applies.

Basically, Chapter V of the GDPR ensures that the protection that the GDPR offers to a natural person with regard to its personal data travels with the data when it leaves the EEA territory. That is why the GDPR restricts the transfer of personal data outside the EEA. Only when one of the *'transfer mechanisms'* listed in articles 45 to 47 of the GDPR is complied with, the transfer of personal data to a third country or international organisation may take place.

➢ According to **article 45 (1) of the GDPR**, a transfer of personal data to a third country or an international organisation may take place where the European Commission (EC) has decided that the third country, a territory or one or more specified sectors within that third country, or the international organisation in question ensures an *adequate level of protection*, i.e. if the EC has made an **adequacy finding** for the country or territory concerned. In that case, the transfer of personal data to that third country or international organisation will not require any specific authorisation.

➢ In the absence of such an adequacy decision, according to **article 46 (1) of the GDPR**, personal data may be transferred to a third country or international organisation only if the controller or processor has provided **appropriate safeguards**, and on the condition that enforceable data subject rights and effective legal remedies for data subjects are available.

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)



These appropriate safeguards may be provided for by:

- A legally binding and enforceable instrument between public authorities or bodies

  This is not an appropriate safeguard when the data exporter (based in the EU) or the data importer (based outside the EU) is a private body or an individual.

- Binding corporate rules (BCR's)

  These are personal data protection policies that serve as internal rules for data transfers within multinational companies. BCR's have to be authorised by the supervisory authority(ies) before any transfer can be performed.

- Standard data protection clauses (adopted by the Commission or adopted by a supervisory authority and approved by the Commission)

  These are 'model contract clauses' that should in their entirety be incorporated into a contract between the data exporter (based in the EU) and the data importer (based outside the EU), before the transfer can be performed. The clauses contain contractual obligations on the data exporter and the data importer as well as rights for the individuals whose personal data is transferred.

- An approved code of conduct (together with commitments of the data importer in the third country to apply the appropriate safeguards)

  This option is newly introduced by the GDPR and no approved codes of conduct are yet in use.

- An approved certification mechanism (together with commitments of the data importer in the third country to apply the appropriate safeguards)

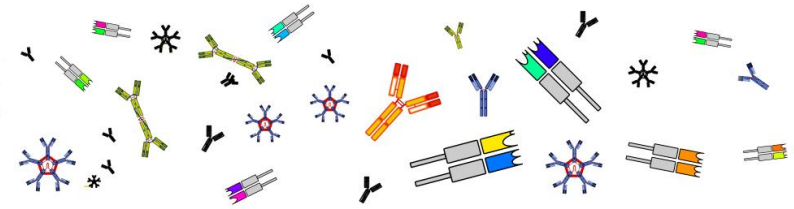
  This option is newly introduced by the GDPR and no approved codes of conduct are yet in use.

- Contractual clauses between the data exporter and the data importer in the third country or international organisation, authorised by the competent supervisory authority.

  These are data protection clauses incorporated in a contract between the data exporter and the data importer that haven been individually authorised by the supervisory authority of the country from which the data are exported.

- Administrative arrangements between public authorities or bodies which include enforceable and effective rights for the individuals whose personal data is transferred, and which have been authorised by a supervisory authority

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)



This is not an appropriate safeguard for restricted transfers between a public and private body. Moreover, this option is newly introduced by the GDPR and no approved administrative arrangements are yet in use.

➢ If, however no adequacy decision or appropriate safeguards as mentioned above are in place, a transfer of personal data to a third country or an international organisation can only take place if one of the **derogations for specific situations** listed in **article 49 of the GDPR** applies.

This is the case when:

- the data subject has explicitly *consented* to the proposed transfer, after having been informed of the possible risks thereof in absence of adequacy decision or appropriate safeguards;

- the transfer is *necessary for the performance of a contract between the data subject and the controller* or the implementation of pre-contractual measures taken at the data subject's request;

- the transfer is *necessary for the conclusion or performance of a contract concluded in the interest of the data subject* between the controller and another natural or legal person;

- the transfer is *necessary for important reasons of public interest*;

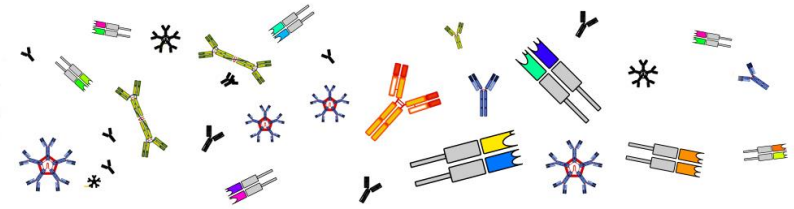
- the transfer is *necessary for the establishment, exercise or defence of legal claims*;

- the transfer is *necessary in order to protect the vital interests of the data subject or of other persons*, where the data subject is physically or legally *incapable of giving consent*;

- the transfer is made from a *register* which according to Union or Member State law is *intended to provide information to the public and which is open to consultation* either by the public in general or by any person who can demonstrate a legitimate interest, but only to the extent that the conditions laid down by Union or Member State law for consultation are fulfilled in the particular case.

Lastly, if the transfer cannot be based on one of the derogations set out above, a transfer may only take place if it *is non-repetitive, concerns only a number of data subjects, is necessary for the purposes of compelling legitimate interests of the controller that are not overridden by the data subject* and the controller has provided *suitable safeguards* with regard to the protection of personal data.

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)



2.3 Other requirements with regard to transfers of personal data

Besides the requirements resulting from Chapter V of the GDPR, it should not be forgotten that other requirements laid down in the GDPR may also apply, given that *a transfer of personal data to a non-EU country or international organisation in itself implies a processing activity*.

As was briefly mentioned above, every processing of personal data should be supported by one of the 6 legal bases listed in the GDPR. Given that a transfer of personal data to a third country or international organisation implies a processing activity, it should not only be justified by one of the transfer mechanisms discussed, but it should also be backed by a legal basis.

Pursuant to **article 6 of the GDPR**, any processing of personal data can only be lawful if and to the extent that one of the following **legal bases** apply:

- The data subject has given *consent* to the processing of his/her personal data for one or more specific purposes;
- The processing is *necessary for the performance of a contract* to which the data subject is party or the processing is necessary in order to take steps at the request of the data subject prior to entering into a contract;
- The processing is *necessary for compliance with a legal obligation* to which the controller is subject;
- The processing is *necessary in order to protect the vital interests* of the data subject or of another natural person;
- 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;
- The processing is *necessary for the purposes of the legitimate interests* pursued by the controller or by a third party.

This legal basis does not apply however when these interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data.

Thus, when a transfer of personal data to a third country or international organisation is envisaged, the controller or processor must not only comply with the specific requirements of Chapter V of the GDPR, but must also be able the show which legal basis in the list of article 6 of the GDPR applies in order for the transfer to be lawful.

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)

Furthermore, if the processing activity involves **'special categories of personal data'**[5], such as genetic data, biometric data or health data, the processing is in principle prohibited.

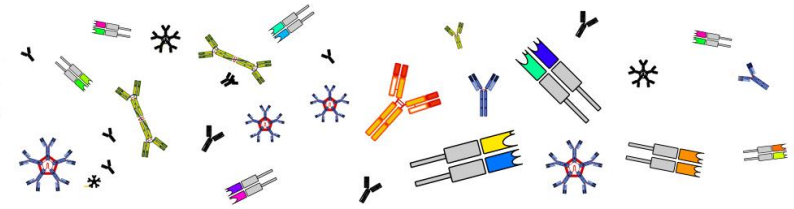
Nevertheless, processing of these special categories of personal data will be allowed, if one of the **exceptions of article 9 GDPR** applies. This is (amongst others) the case when the processing of the personal data is *necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes*.

In that case article 89(1) of the GDPR must be complied with, which requires that **appropriate safeguards** be put in place for the rights and freedoms of the data subject. These safeguards shall ensure that **technical and organisational measures** are in place, in particular in order to ensure respect of the principle of data minimisation (e.g. by pseudonymising the personal data).

Moreover, if special categories of personal data are processed for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, the processing for these purposes must have a **basis in Union or Member State law**.

[5] Special categories of data considered in the GDPR are: personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership as well as genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health and data concerning a natural person's sex life or sexual orientation.

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)



3 Analysis of the data processing activities in iReceptor Plus

In order to be able to assess the data processing activities performed in the iReceptor Plus action in light of the GPDR, it is necessary to first discover which data is processed in the course of the action (section 3.1), where it is stored and where it might be transferred to (section 3.2).

3.1 Data processed in iReceptor Plus

It is useful to recall that iReceptor Plus will be developed as a freely and openly available software platform, which operates on three levels. First of all (1), the iReceptor Plus platform will give access to publicly available AIRR-seq data in the AIRR Data Commons[6] ('public AIRR-seq data'[7]). On a second level (2), the iReceptor Plus platform will provide an intermediate level of sharing of non-publicly available AIRR-seq data that can only be accessed if common consent structures or reciprocal data transfer agreements (DTA) are put in place ('controlled AIRR-seq data'[8]). Finally (3), iReceptor Plus will provide the ability to integrate public AIRR-seq data with non-public non-AIRR seq data (such as information extracted from individual electronic health records), without exposing the non-public non-AIRR seq data.

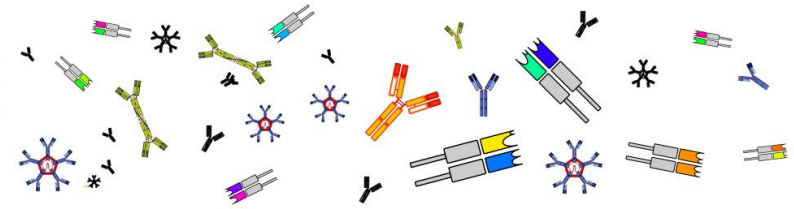
This way, the iReceptor Plus platform will not only provide access to large amounts of public data but will also facilitate the comparison and integration of non-public data with public data.

- The **public AIRR-seq data** (including metadata) are data that were originally used to perform scientific research, and which were afterwards deposited in a repository as a pre-requisite to the publication of the research paper the data support.

  In the context of iReceptor Plus, the individual repositories include both AIRR-seq data that come directly from research studies performed by the scientific entity curating the repository as well as AIRR-seq data that were taken from other large repositories

[6] The AIRR Data Commons is the network of all distributed AIRR-seq data repositories.
[7] Public AIRR-seq data are AIRR-sequences, including metadata that are made available to all through a public repository.
[8] Controlled AIRR-seq data are AIRR-sequences, including metadata that are not made available to all through a public repository, but can only be accessed if access is authorized by the data owner through a reciprocal data transfer agreement or common consent structure concluded with the person or entity envisaging to access the data. In essence, the data owner 'controls' the access to the data.

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)



such as SRA[9] or ENA[10].

Every deposit of AIRR-seq data into a repository is always performed under the supervision of and with the authorisation of the ethics committee responsible for the research data concerned. Each ethics committee is responsible for following ethical guidelines and complying with the GDPR and national data protection laws. Given that most national laws foresee an obligation of anonymisation before research data can be published, the AIRR-seq data deposited in the public repositories should be considered anonymous.

- The **controlled AIRR-seq data** (including metadata) are data that do not originate from publicly available repositories but come from research entities that have themselves performed the research concerned. Since those research entities were in contact with the people from whom the samples underlying the data were taken, they might be able to retrace the people to whom the data relate, unless anonymisation techniques were already applied to the data. This leads to the conclusion that some of the controlled AIRR-seq data may be non-anonymous data.

- The **non-public non-AIRR seq data** that can be integrated with public AIRR-seq data via the iReceptor Plus platform might be anonymous or non-anonymous data depending on the question if the entity curating these data has applied anonymisation techniques to the data. However, when integrating these non-public non-AIRR seq data with public AIRR-seq data, the non-public non-AIRR seq data will always remain with entity requesting the integration with the public AIRR-seq data and will never be shared with other parties. Consequently, this integration capability of the iReceptor Plus platform will never lead to a transfer of personal data from an EEA-country to a non-EEA country and does not raise concerns in the context of this deliverable.

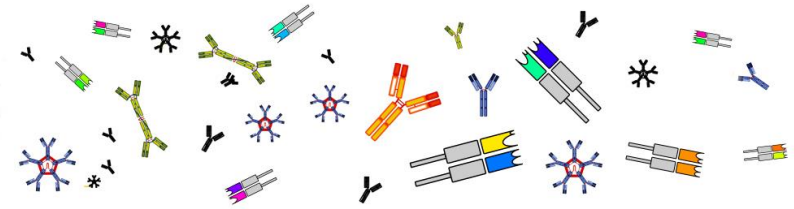
During the **initial development phase** of the iReceptor Plus project, **only public AIRR-seq data curated in the distributed repositories will be processed.** Given that access to controlled AIRR-seq data requires common consent structures or reciprocal transfer agreements, these data will not be part of the initial development phase (M1-M24) of iReceptor Plus.

After the initial development phase, when the security and access control measures required to enable researchers to manage data according to their data transfer agreements and consent structures, **controlled AIRR-seq data will also be processed**. Knowing that these data may in some cases be non-anonymous data and thus personal data, as this later stage of the project, the

[9] The Sequence Read Archive (SRA) is a bioinformatics database that provides a public repository for DNA sequencing data, especially the 'short reads' generated by high-throughput sequencing, which are typically less than 1000 base pairs in length. .

[10] The European Nucleotide Archive (ENA) is a repository providing free and unrestricted access to annotated DNA and RNA sequences.

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)



requirements for transfers of personal data (as set out above) will have to be complied with. This will be confirmed in future deliverables that deal with re-assessing the need for GDPR-compliance and data management compliance (such as D3.4, M24).

3.2 Storage and transfer of data in iReceptor Plus

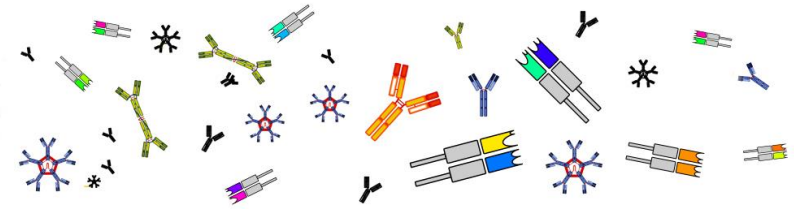
iReceptor Plus will be designed as a network of federated repositories that facilitates data queries and advances analysis through a centralized web portal. This distributed, federated approach is central to the iReceptor Plus concept as it allows each institution to maintain control over their data and stay compliant with local legislation.

This approach also entails that the data that are processed in the course of iReceptor Plus, will be stored in the individual repositories that underly the iReceptor Plus network. Below are listed the beneficiaries of whom the repositories will probably be involved in iReceptor Plus, as well as the storage location.

Beneficiary		**Storage location**
Bar Ilan University	BIU	Israel
Simon Fraser University	SFU	Canada
Sorbonne University	SORBONNE	France
University of Toronto	UTORONTO	Canada
German Cancer Research Center	DKFZ	Germany
University of Haifa	UH	Israel
University of Oslo	UiO	Norway
Medgenome Inc.	MedGenome	United States of America
10X Genomics	10X	United States of America
Clalit health services	Rabin MC	Israel
University of Texas System	UTSW	United States
Oslo University Hospital	OUS	Norway

From the above table it follows that data processing activities throughout the development of the iReceptor Plus platform may cross EEA-borders, in a way that data originating from within the EEA will be transferred to organizations located outside of the EEA.

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)



4 Assessment of the data processing activities in light of the GDPR

Now that the data processing activities of iReceptor Plus have been analysed and the requirements for the transfer of personal data to third countries or international organisations have been clearly set out, it should be assessed if the datasets exchanged in the context of iReceptor Plus are governed by the GDPR and Chapter V of the GDPR in particular.

Section 3 of this deliverable explained that only public AIRR-seq data will be processed in the initial development phase of the action (M1-M24). Section 2 of this deliverable clarified that the GDPR only applies to the processing of *personal* data. Consequently, it should be assessed if the public AIRR-seq data processed in the action are to be considered as personal data in the sense of the GDPR, or anonymous data.

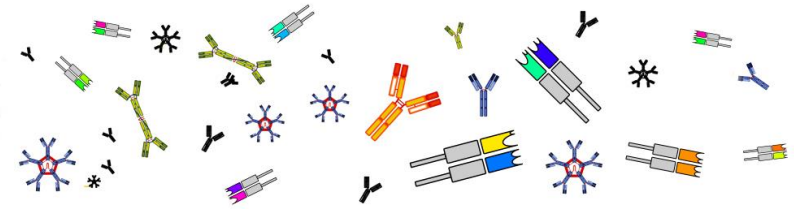
Regarding anonymous data, the GDPR states that 'the *principles of data protection should not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that a data subject is no longer identifiable'.* Therefore, the GDPR does not concern the processing of such anonymous information, including for statistical or research purposes.

This raises the question what should be considered as non-identifiable (and thus anonymous) data in the sense of the GDPR. The answer to this question can be found in recital 26 of the GDPR which reads as follows: *'To determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly. To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments.'*

This means that for data to be processed anonymously, (1) it must be stripped of any information that could lead to identification, either directly or indirectly – for instance, by combining the data with other available information; (2) that this process of de-identification must be irreversible; and (3) that de-identification assessment should focus on outcomes, rather than means or procedures.[11]

[11] See Working Party 29, Opinion 05/2014 on Anonymisation Techniques, 10 April 2014, p. 5: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf and M. SHABANI and L. MARELLI, "Re-identifiabilit of genomic data and the GDPR Assessing the re-identifiability of genomic data in light of the EU General Data Protection Regulationn", EMBO REPORTS, doi: 10.15252/embr.201948316.

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)



In literature, there is an ongoing discussion on the question if genomic data (such as the AIRR-seq data in iReceptor Plus) can be truly irreversibly de-identified. Can genetic data ever be considered as non-personal data for the purpose of the GDPR? Up until now, this question has not yet been decided upon.

Anyhow, according to literature[12], the assessment if certain data is to be considered as anonymous data can be made by assessing the likelihood of re-identification of these data. Important factors to assess this likelihood of re-identification are: the peculiar characteristics of a specific genetic dataset (such as the type of data, sample size, rareness of the genetic variant considered), the institutional setting in which processing occurs (health care, research, consumer genomics, forensics), the stage of processing, the availability of cross-referenceable datasets, the availability of resources for re-identification, the mitigation strategies adopted. The question if certain data are irreversibly de-identified and thus anonymous, is therefore always to be answered on a case-by-case basis. This is also shown by multiple studies that have been successful in re-identifying alleged anonymous genomic data.

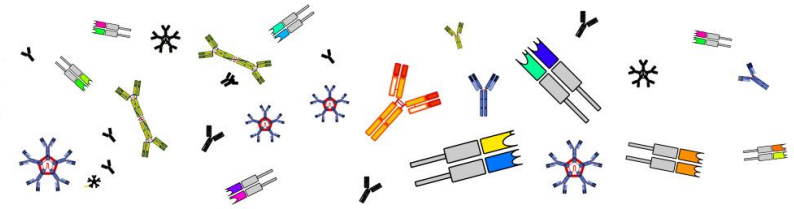
Nevertheless, it should not be forgotten that according to the GDPR the assessment of the identifiability of a person through certain data must be made considering all the means *reasonably likely to be used* to identify the natural person directly or indirectly.

Given that the AIRR-seq data that will be processed throughout the first phase (M1-M24) of the iReceptor Plus research action consist only of public data that have undergone anonymisation techniques, to our understanding, these data are non-personal data. This is because re-identifying these data would require unreasonable means in terms of costs, time and technology, making re-identification highly unlikely. Moreover, the AIRR-seq data processed in the iReceptor Plus action were published together with the studies they support, as is common practice in scientific research. Finally, in almost all European countries researchers who publish such data without ensuring adequate anonymisation would violate their national data protection laws. National legislators have included in these laws the principles of the Council of Europe Recommendations, in particular of Recommendation No R(97)5 on the protection of medical data. Paragraph 12.5 of this document explicitly states that *personal data used for scientific research may not be published in a form which enables the data subjects to be identified*, unless they have given their consent for the publication and publication is permitted by domestic law.

For these reasons, to our current knowledge, the datasets exchanged between participating repositories in the first phase of the iReceptor Plus action cannot be qualified as personal data.

[12] See for example: M. SHABANI and L. MARELLI, "Re-identifiabilit of genomic data and the GDPR Assessing the re-identifiability of genomic data in light of the EU General Data Protection Regulationn", EMBO REPORTS, doi: 10.15252/embr.201948316.

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)



Consequently, the transfer of public AIRR-seq data to a non-EEA country or international organisation throughout the initial phase of the action (M1-M24) does not require a transfer mechanism in the sense of Chapter V of the GDPR.

However, to mitigate any remaining risk of re-identification, which would render the public AIRR-seq data personal (and make the GDPR applicable), the beneficiaries will consider adherence to the guidance of sectoral codes of conduct or international guidelines by professional societies on the concrete implementation of the GDPR with regard to the processing of genomic data in scientific research, such as for example the BBMRI-ERIC GDPR Code of Conduct for health research.

In any case, the consortium is aware that the assessment of the identifiability of the genetic data concerned is a dynamic exercise which cannot be decided upon once and for all. The beneficiaries therefore commit to periodically re-assess the risk of re-identification at every stage of processing. This aligns with the risk-based approach to data protection adopted by the GDPR through 'accountability'[13] and 'data protection by design and by default'[14] principles (art. 5(2), 24 and 25 GDPR).

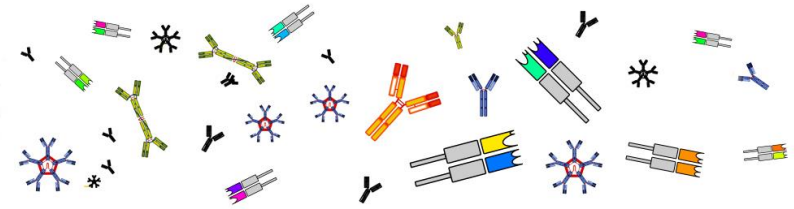
Should it at a certain stage become clear that the processing activities carried out in the research and innovation action do involve personal data in the sense of the GDPR (for example, as of M24, when controlled AIRR-seq data will become accessible), then the beneficiaries ascertain that the requirements of the GDPR, including those of Chapter V, will be duly complied with. In that case, in accordance with Chapter V of the GDPR, in our view, it would be best to use the standard contractual clauses adopted by the EC as a transfer mechanism between the EEA and non-EEA repositories.

That is why the technically skilled beneficiaries of the action will monitor the data processing activities performed in the action and will report on any changes thereof to the legal and ethics partners involved in iReceptor Plus. This will allow iReceptor Plus to re-assess its current compliance whenever necessary, so as to remain GDPR-compliant throughout the entire project. Such re-assessment will, in our opinion, in any case prove necessary as of month twenty-four of the project.

[13] The accountability principle requires controllers to put in place appropriate technical and organisational measures and be able to demonstrate what they did and its effectiveness when requested.

[14] The principles of data protection by design and default requires controllers to integrate data protection into the processing activities and business practices from the design stage right through the lifecycle. The principles are about considering data protection and privacy issues upfront in everything you do. The principles form part of the focus on accountability.

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)



5 Conclusion

The analysis performed in this deliverable has shown that the AIRR-seq data exchanged between participating repositories in the course of the initial phase of the iReceptor Plus action can be legally qualified as anonymised data. Knowing that anonymous data are not considered to be personal data in the sense of the GDPR, the GDPR is not applicable to the iReceptor Plus core processing activities as currently performed. Nevertheless, to mitigate any risk that in this first phase of the action personal data would nonetheless be processed or that anonymous data could potentially be re-identified, the beneficiaries will (1) consider adhering to the BBMRI-ERIC GDPR Code of Conduct for health research and will (2) monitor and periodically re-assess the identifiability of the data that are processed during the initial phase of the iReceptor Plus action as well as in the later stages of the project, with the aim of ensuring proper compliance to the highest standards of data protection, as enshrined in the GDPR.

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)