DELIVERABLE 8.2

LEGAL, ETHICS, PRIVACY & SECURITY FRAMEWORK

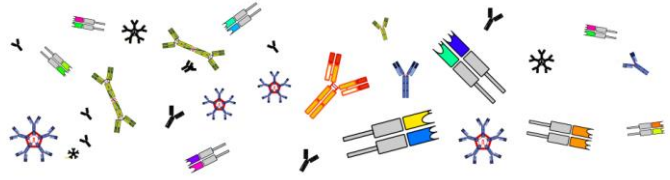
1

WORK PACKAGE NUMBER: WP8

WORK PACKAGE TITLE: INNOVATION MANAGEMENT, EXPLOITATION & BUSINESS PLANNING

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

Deliverable Information	
Work package number	WP8
Work package title	Innovation Management, Exploitation & Business Planning
Deliverable number	D8.2
Deliverable title	Legal, ethics, privacy & security framework
Description	Legal, ethics, privacy & security framework, reviewing and structuring key challenges of this domain deemed relevant for further project work
Lead beneficiary	time.lex
Lead Author(s)	Jos Dumortier
Contributor(s)	Liesa Boghaert
Revision number	3
Revision Date	27 August, 2019
Status (Final (F), Draft (D), Revised Draft (RV))	F

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)

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))	PU

Document History			
Revision	**Date**	**Modification**	**Author**
1	21/08/2019	First draft	Jos Dumortier
2	21/08/2019	Review/edit	Brian Corrie
3	22/08/2019	Review	Tobias Hinz

Approvals				
	Name	**Organisation**	**Date**	**Signature (initials)**
Coordinator	Prof. Gur Yaari	Bar Ilan University	30.08.2019	GY
WP Leaders	Tobias Hinz	Ascora	30.08.2019	TH

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)

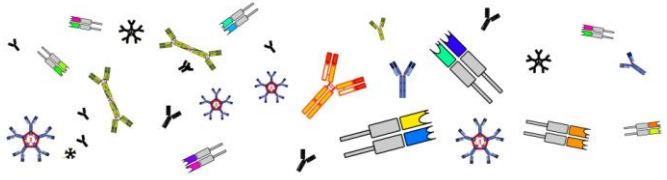
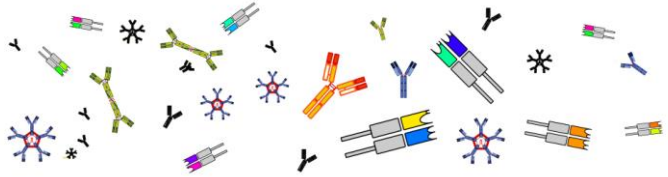


Table of Contents

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)

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 'Legal, ethics, privacy & security framework' (D8.2) defines the initial legal, ethical, privacy and security requirements for iReceptor Plus at a high level. It is intended to support the alignment of iReceptor Plus with the highest standards of legal, ethical, privacy and security compliance existing in both EU and other participating countries.

The purpose of this document is to develop a framework for identifying the core legal, ethical, privacy & security requirements for the execution of iReceptor Plus, without at this stage delving deep into operational implementation. It is aimed at reviewing and structuring key challenges and developing a relevant project policy towards resolving them, across 10 specific domains:

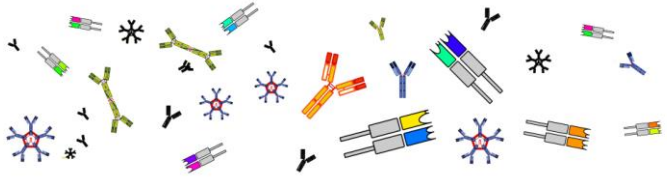
1. Anonymisation and pseudonymisation
2. Informed consent, fairness and transparency
3. Lawfulness and further processing of personal data
4. Continuous risk assessment
5. Data protection by design and default
6. Transfer of personal data
7. Data security, deletion and archiving
8. Big data and artificial intelligence
9. Data ownership and IP rights
10. National compliance

In each of these domains, the high-level issue is reflected upon in light of iReceptor Plus and is then linked to several requirements for the project. In this manner, the deliverable can act as an elementary legal and ethical guide for the activities performed in iReceptor Plus.

Task 8.2 (Legal and ethics framework and process) is continuous throughout the project: while this first version of the deliverable aims to outline and structure key challenges deemed relevant for further project work, the legal, ethical, privacy and security requirements set forth in this deliverable will be further addressed in future iterations (M24 and M36) as well as in other deliverables, such as D3.4 (Monitoring and GDPR aspects) and D3.5 (Recommendations and proposals for new regulatory regimes).

Hence, this 'Legal, ethics, privacy & security framework' should not be seen as a fully mature and exhaustive document, but rather as a first overview of essential challenges and compliance strategies that will be continued and detailed throughout the project, based on additional input from iReceptor Plus partners and the further identification of relevant factors and regulatory processes at the EU level and in each of the countries participating.

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

Legal and ethical issues are core aspects for a sustainable implementation of the iReceptor Plus platform and are vital to guarantee a successful use of the platform by end-users.

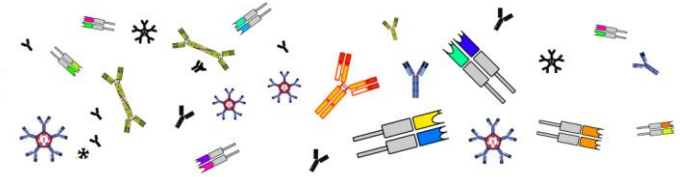
As a part of Work Package 8 (Innovation Management, Exploitation & Business Planning), Task 8.2 - 'Legal and ethics framework and process' is dedicated to ensuring that iReceptor Plus is executed in accordance with the highest standards of legal, ethical, privacy and security compliance existing in both EU and other participating countries.

The purpose of this deliverable 8.2 'Legal, ethics, privacy & security framework' is to develop a framework for identifying the core legal, ethical, privacy & security requirements for the execution of iReceptor Plus, without at this stage delving deep into operational implementation. It is aimed at reviewing and structuring key challenges in iReceptor Plus and developing a relevant project policy towards resolving them.

Methodologically, this will be done by analysing ten domains relevant to iReceptor Plus from a legal, ethical, privacy and security perspective and drawing requirements pertaining to these domains from applicable international, EU and national legislation, soft law, policy and guidance documents. Among these documents are the United Nations International Declaration on Human Genetic Data[1], the OECD Guidelines on Human Biobanks an Genetic Research Databases[2], the EU General Data Protection Regulation (GDPR)[3], the Ethics Guidelines for Trustworthy AI of the High-Level Expert Group on Artificial Intelligence[4], the Canadian Digital Charter[5], the Council of Europe Guidelines on the protection of individuals with regard to the processing of personal data in a world of Big Data[6], the European Commission's guidance on ethics and data protection[7] etc.

[1] Accessible at: http://www.unesco.org/new/en/social-and-human-sciences/themes/bioethics/human-genetic-data/.
[2] Accessible at: http://www.oecd.org/sti/emerging-tech/guidelines-for-human-biobanks-and-genetic-research-databases.htm.
[3] Regulation (EU) 2016 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), accessible at: https://eur-lex.europa.eu/legal-content/NL/TXT/?uri=CELEX%3A32016R0679.
[4] Accessible at: https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai.
[5] Accessible at: https://www.ic.gc.ca/eic/site/062.nsf/eng/h_00108.html.
[6] Accessible at: https://rm.coe.int/16806ebe7a.
[7] Accessible at: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/ethics/h2020_hi_ethics-data-protection_en.pdf.

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)



Task 8.2 is continuous throughout the project: while this first version of the deliverable aims to outline and structure key challenges deemed relevant for further project work, the legal, ethical, privacy and security requirements set forth in this deliverable will be further addressed in future iterations (M24 and M36) as well as in other deliverables, such as D3.3 (Monitoring and GDPR aspects) and D3.4 (Recommendations and proposals for new regulatory regimes).

Hence, this legal, ethics, privacy & security framework should not be seen as a fully mature and exhaustive document, but rather as a first overview of essential challenges, requirements and compliance strategies that will be continued throughout the project, based on additional input from iReceptor Plus partners and the further identification of relevant factors and regulatory processes at the EU level and in each of the countries participating.

In what follows, 10 key domains that pose challenges to iReceptor Plus will be discussed, along with the resulting, legal, ethical, privacy and security requirements of the project. As explained, at this stage, the challenges and requirements are developed at a high level. During the remainder of the project, these will be further developed and implemented in a way that ensures that they meet both EU and relevant national (implementing) laws.

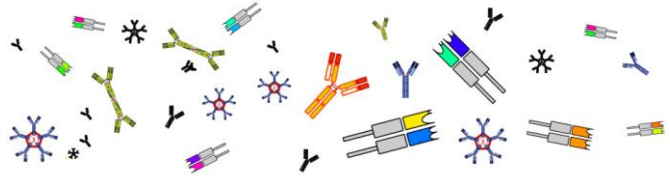
1.2 Outline of iReceptor Plus objectives and relating concerns

Prior to examining the high-level legal, ethical, privacy and security challenges and compliance requirements in the following section of this deliverable, it is worth briefly recalling the objectives pursued in iReceptor Plus.

The iReceptor Plus project essentially intends to lower the barrier to share, access and analyse large sets of Adaptive Immune Receptor Repertoire sequencing data (AIRR-seq data) from around the world and to ease the availability of these AIRR-seq data to academia, industry and clinical partners. This increased availability of AIRR-seq data will advance the understanding of immune responses and may lead to the discovery of biomedical interventions (such as vaccines and other immunotherapies) that manipulate the adaptive immune system. Such advancements will enable improved personalized medicine and immunotherapy in cancer, inflammatory an autoimmune diseases, allergies and infectious diseases.

To this aim, iReceptor Plus will create a distributed network of repositories containing both AIRR-seq data and non-AIRR-seq data (such as clinical data, biological data, sample metadata, receptor reactivity data…), to allow for the analysis of global interactions within the immune system and its environment. Although the distributed nature of the network implies that each research institution will maintain control over its own data and will be responsible for remaining compliant with local legislation, the distributed repositories will be interoperable, and the AIRR-seq 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)



within them will be reusable, given that all the repositories involved will store their data using a common standard (the MiAIRR standard[8]). Moreover, iReceptor Plus will provide a turnkey AIRR-seq repository solution for those researchers that do not have their own system.

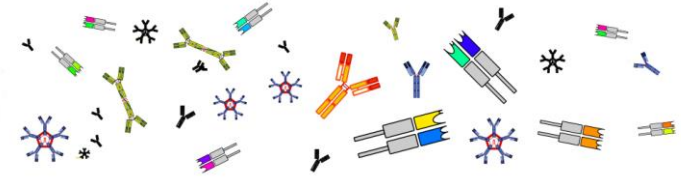
As matters stand, iReceptor Plus will include repositories based in the EU (France, Germany), Israel, Norway, Canada (British Columbia and Ontario) and the United States of America (California and Texas). It is however envisaged that other repositories, from non-beneficiaries in this project will also be participating in the future. Between all of the repositories, data will be exchanged using a web-based Application Programming Interface (API) that is also based on the MiAIRR standard and that is implemented as the iReceptor REST API.

Furthermore, iReceptor Plus will provide a data exploration, aggregation and analysis tool for researchers in the form of the iReceptor Plus Scientific Gateway. This Scientific Gateway will allow researchers to pose complex queries about AIRR-seq data, their metadata and annotated sequence data. The Gateway then, on behalf of the end-user, will send the query to each of the repositories (using the REST API), will federate the results from each repository and present these federated results to the end-user. But the Gateway will go even further, as it will be able to stage federated data resulting from a query to an advanced analysis tool that uses computational methods on the aggregated data, such as relational datamining algorithms and deep learning techniques (AI), to facilitate complex analysis of the federated data and integrate it with other types of human health and genomic data.

iReceptor Plus as a research project therefore combines many potential legal, ethical, privacy and security concerns. These concerns relate to the fundamental characteristics of the project. First of all, iReceptor Plus will build a platform to integrate distributed repositories of AIRR-seq and other data. This *distributed, federated approach* implies that large amounts of data will be shared cross-borders. Given that the repositories currently planned to take part in iReceptor Plus are based in and outside the EU, both EU legislation and national legislation applicable to the participating repositories will have to be taken into account. Moreover, the iReceptor Plus platform will allow for the querying, aggregation and analysis of *AIRR-seq data and non-AIRR-seq data* (such as clinical and biological data). AIRR-seq data are genetic data and should therefore be considered as particularly sensitive, in a way that they raise significant data protection and security challenges. These concerns are intensified by the fact that AIRR-seq data will be complemented with non-AIRR seq data that allow for the analysis of the global interactions with the immune system and its environment. The integration of these data can be more highly privacy sensitive and will thus require higher privacy and security measures. Furthermore,

[8] This is an AIRR Community-endorsed standard that describes the 'Minimal Information' for repertoire metadata and sequence annotation data in studies that utilize AIRR-seq data. The AIRR Community is a research-driven group that is organizing and coordinating stakeholders in the use of next-generation sequencing (NGS) technologies to study antibody/B-cell and T-cell receptor repertoires.

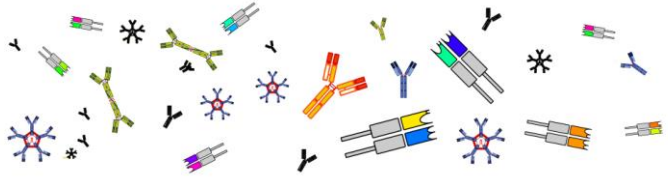
This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)



iReceptor Plus aims to apply a *big data analytics approach*, in which data sets from multiple repositories are combined. These data exploration, aggregation and analysis capabilities increase the importance of appropriate anonymisation or pseudonymisation strategies to mitigate privacy and security challenges. In this context, the implementation of artificial intelligence (AI), including deep learning techniques, also raises additional ethics questions. Lastly, the iReceptor Plus platform will be available as a *free and open license software* and will offer a *turnkey repository solution* to researchers. This means that questions of intellectual property rights and data ownership must be addressed as well.

Consequently, it is clear that the main challenges in iReceptor Plus relate to privacy and data protection, as well as the security and ethical requirements that go along with them. Taking into account these considerations, a strict, continuously re-evaluated and continuously developed risk approach and risk assessment is needed, in which each compliance measure is robust, monitored and enforced. In the section below, a first draft of such approach will be presented.

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 Legal, ethical, privacy and security challenges in iReceptor Plus

As noted above, the purpose of this deliverable is to develop a framework for identifying the core legal, ethical, privacy & security requirements for the execution of iReceptor Plus. In the sections below, we will define and discuss at a high level the key challenges that are deemed relevant for further project work. In future iterations of this deliverable as well as in other deliverables (D3.4 and D3.5), the requirements resulting from these challenges will be further developed and compliance will be shown.

2.1 Anonymisation and pseudonymisation

2.1.1 Meaning and relevance to iReceptor Plus:

One of the best ways to mitigate data protection and privacy harms and to reduce security risks, is to keep the processing of personal data[9] at a minimum. One of the fundamental principles of EU data protection law, as found in the GDPR, states that personal data shall be adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed (the data minimisation principle). This principle can be found in many other data protection laws, such as the 'Fair information principles'[10] of the Canadian PIPEDA[11]. Principle 4 states that: 'The collection of personal information must be limited to that which is needed for the purposes identified by the organization.'. In the same sense, the Californian HIPAA[12] Minimum Necessary Rule states that: 'A covered entity must make reasonable efforts to limit the scope of the PHI it uses, discloses or requests to the minimum necessary to accomplish the intended purpose of the use, disclosure or request.'.

Within iReceptor Plus, the data minimisation principle implies that the data in the distributed repositories shared through the iReceptor Plus network should to the greatest extent possible be anonymized, so that they no longer relate to identifiable persons. Data that no longer relate to identifiable persons, such as aggregate data and statistical data, or data that have otherwise been rendered anonymous in a way that the person underlying the data cannot be re-identified, are

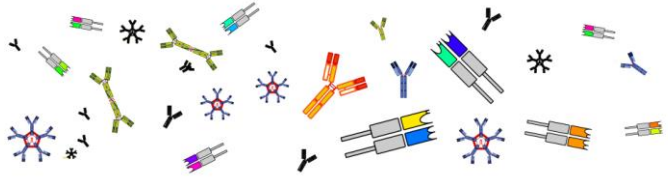
[9] Please note that for the purpose of this deliverable no distinction will be made between the EU-law term 'personal data' and the terms 'personal information' and 'information', which are used in other jurisdictions (such as the USA, Canada and Israel) and which may comprise different kinds of data. In what follows, the term 'personal data' will be used to cover all these data.

[10] These principles form the ground rules for the collection, use and disclosure of personal information, as well as for providing access to personal information. See: https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-personal-information-protection-and-electronic-documents-act-pipeda/p_principle/.

[11] PIPEDA stands for 'Personal Information Protection and Electronic Documents Act'. The Act can be found at: https://laws-lois.justice.gc.ca/eng/acts/P-8.6/.

[12] Health Insurance Portability and Accountability Act.

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)



not personal data and therefore raise less ethical, privacy and data protection concerns. The same applies for non-EU data protection laws, none of which apply to non-personal information.

However, it may prove difficult to create fully anonymous datasets that still include the granular information needed for research purposes. Some research requires a link between the research subjects and their personal data to be retained in order for valuable results to be achieved. If that is the case, the data concerned should be pseudonymised, i.e. processed only in such a manner that the personal data can no longer be attributed to a specific person without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable person. These pseudonymised data should nevertheless be considered as personal information that require compliance with all applicable data protection laws.

2.1.2 Resulting requirements:

- Data processed in iReceptor Plus should be anonymized to the greatest extent possible. Sharing public AIRR-seq data[13] through the iReceptor Plus network should thus be conditional upon strict anonymisation, and subject to local ethics approval.

- Data that cannot be fully anonymised without losing the granular information needed for the research purposes envisaged by iReceptor Plus should only be shared through the iReceptor Plus platform if it is sufficiently pseudonymised. In that case, all applicable requirements with regard to the protection of personal data will have to be complied with. These requirements may differ depending on the applicable laws.

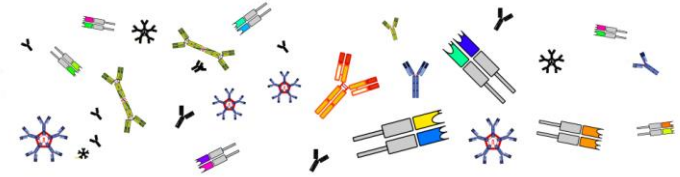
2.2 Informed consent, fairness and transparency

2.2.1 Meaning and relevance to iReceptor Plus:

Informed consent is one of the key pillars of research ethics. Human participation in research projects requires the voluntary, free and informed consent of those who contribute their time, insights, effort and data for the use of researchers. From an ethical point of view, whenever personal data is collected from research participants, they must give their consent to the processing of that data after having been informed about the subject and objectives of the research, the risks involved etc. This obligation relates to the respect that researchers owe to the right to integrity and autonomy of research participants. However, strictly speaking, according to

[13] Public AIRR-seq data are publicly available AIRR-seq data in the AIRR Data Commons. These 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.

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)



EU law, personal data may, under certain circumstances, be processed for research purposes without the consent of the research participant.[14]

If the processing of personal data is however based on the consent of the research participant, this consent and the information given to the research participants must, in accordance with the principles of fair and transparent data processing, cover all the data processing activities related to their participation in the research. Nevertheless, given that at the time of data collection, fully identifying the purpose of personal data processing for scientific research purposes is often not possible, it is sufficient if consent is given for certain areas of scientific research, if this is in conformity with recognised ethical standards for scientific research.[15]

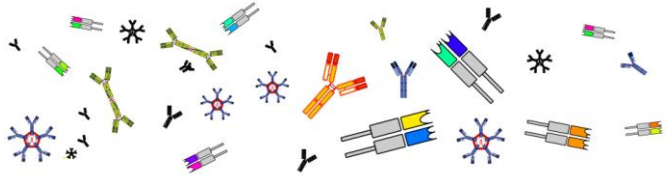
Under EU law, consent requires a clear affirmative act establishing a freely given, specific, informed and unambiguous indication of a person's agreement to the processing of their personal data.[16] Consent is 'informed' if the person concerned is provided with detailed information about the envisaged data processing in an intelligible and easily accessible form, using clear and plain language.[17] This information should include (among others) the purposes of the processing, the storage period and the sharing of data with research partners or organisations outside of the EU.

iReceptor Plus as a research project does not involve any human participation. Nevertheless, the research project aims at developing a network of distributed repositories containing AIRR-seq data coming from medical treatment and research or clinical studies that did involve human participation. As such, it is important that these patients and research participants have consented to the processing of their personal data in the original medical treatment or research settings (unless another legal basis for processing the personal data exists). Under EU law, when consent is used as a legal basis for processing personal data, it needs to be documented in records showing the informed consent procedure, the consent forms provided to patients and research participants and the consent given. This relates to the ethical principle of accountability.

These requirements regarding consent of research participants are however for the data stewards curating the repositories to fulfil. Prior to sharing non-anonymous, personal data through the iReceptor Plus platform, every data steward of an individual repository should verify that an appropriate legal basis for processing personal data is used, such as the consent of the research participants. iReceptor Plus as a research project is mainly involved with software development and will thus not engage in managing or documenting the consent procedures of all the repositories involved.

[14] See section 2.3.
[15] Recital 33 GDPR.
[16] Recital 32 GDPR.
[17] Article 12 and recital 42 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)



In fact, iReceptor Plus will merely facilitate and support the sharing of data by the individual repositories, by offering them an easy way to manage the authentication and authorization of users and restrict access to datasets where needed in accordance with the given consent.

Although iReceptor Plus will require the data stewards of the individual repositories to confirm that they have a legal basis for sharing any personal data, the actual responsibility of meeting ethics and legal requirements (such as consent requirements) remain with the data steward. This includes the decision to upload that data to any repository using the iReceptor Plus repository software.

Institutions that make use of the iReceptor Plus repository software to host a repository (including iReceptor Plus partner institutions) are thus responsible for determining if the iReceptor Plus repository software is appropriate for protecting data stored in their repository at a level that is appropriate for the ethics and privacy constraints applicable, and if the personal data can be shared through the iReceptor Plus platform in a way that does not run counter to the consent given by research participants.

2.2.2 Resulting requirements:

- iReceptor Plus will require data stewards to confirm that personal data are only shared through the iReceptor Plus platform in accordance with the consent acquired from research participants.

2.3 Lawfulness and further processing of personal data

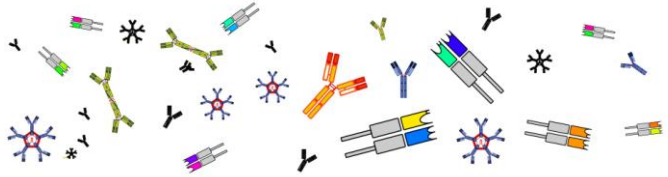
2.3.1 Meaning and relevance to iReceptor Plus:

iReceptor Plus aims to create a network of federated repositories that will allow researchers to query and analyse AIRR-seq data and other information curated in the distributed repositories taking part in the iReceptor Plus project. To the extent that personal data will be accessible through the iReceptor Plus platform, this data sharing capability of the iReceptor Plus platform should be considered as a processing activity regarding the personal data queried or analysed. Moreover, this processing activity constitutes the 'further processing' of the personal data collected and processed in the original medical treatment or research setting.

Although not all data protection laws are as strict, under the EU GDPR all processing activities and further processing activities relating to personal data require a legal basis, as well as other conditions to be fulfilled for the processing to be lawful.

In iReceptor Plus, the data collected and processed by the individual repositories and subsequently shared through the iReceptor Plus platform may involve genetic or health 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)



Pursuant to the GDPR, genetic and health data should be considered as a special category of data for which processing is prohibited.[18]

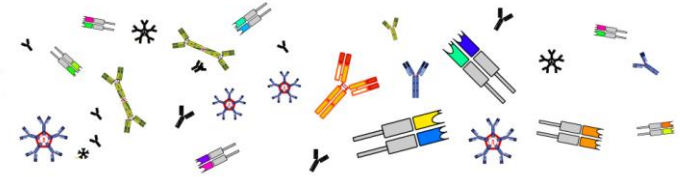
Processing of these **sensitive data** may however still be allowed under certain conditions[19], such as when the person to whom the data relate has consented to the processing or when the processing is necessary to protect the vital interests of the person concerned or another person and where he/she is physically or legally incapable of giving consent (e.g. in a medical treatment setting). Processing of sensitive data is also allowed when it is necessary for scientific research purposes, if that processing has a basis in Union or Member State law and if appropriate safeguards are put in place for the rights and freedoms of the persons concerned.[20] These safeguards should ensure that technical and organisational measures are in place in particular in order to ensure respect for the principle of data minimization (e.g. pseudonymization, implementing an Information Security Management System, cryptographic protection during storage and transfer, authentication, authorization, physical and logical access to data, access logging…).

Moreover, in addition to these conditions for processing sensitive data, a **legal basis for the processing** of these personal data is also required.[21] This legal basis for processing generally can be (a) the consent of the person concerned, (b) the necessity for the performance of a contract to which the person concerned is a party or pre-contractual arrangements, (c) the necessity to comply with a legal obligation to which the controller is subject, (d) the necessity to protect the vital interest of the person concerned or of another natural person, (e) the necessity for the performance of a task carried out in the public interest or the exercise of official authority vested in the controller, (f) the necessity for the purposes of the legitimate interests pursued by the controller or a third party, except where these are overridden by the interests or fundamental rights and freedoms of the person concerned which require protection of personal data, in particular where the person concerned is a child.

In the context of iReceptor Plus, the data included in the distributed repositories either come from research or medical treatment settings. When personal data were originally collected and processed in a medical treatment setting, these processing activities can be based on several legal grounds, such as the consent of the patient (a), the necessity for the performance of a medical contract with the patient (b), the necessity to protect the vital interest of the person concerned (d)… When personal data were originally collected and processed for research purposes, these processing activities can be based not only on the consent of the research participant (a), but also based on the necessity for the purposes of the legitimate interests pursued by the controller or a third party (f). Although research in itself is not considered as a

[18] Article 9 (1) GDPR.
[19] Article 9 (2) GDPR.
[20] Article 9 (2) (j) j° 89(1) GDPR.
[21] Article 6 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)



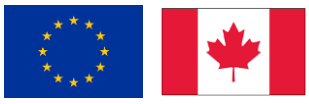
lawful basis for processing under the GDPR, it may qualify as a legitimate interest of the controller. This was confirmed by the Article 29 Working Party[22], which stated that processing for research purposes may constitute a legitimate interest.[23] This however requires a balancing test, in which it is assessed if the legitimate interests of the controller or a third party are not overridden by the interests or fundamental rights and freedoms of the research participant, which require protection of personal data. This assessment should take into account the reasonable expectations of the research participant based on its relationship with the researcher.[24]

As noted above, the sharing of AIRR-seq data by the individual repositories through the iReceptor Plus platform should be considered as a **further processing** for research purposes. According to the GDPR, further processing of personal data is only allowed when the purposes for further processing are compatible with the purposes for which the personal data were collected.[25] Further processing for scientific research purposes is however not considered incompatible with the initial purposes, to the extent that the appropriate safeguards mentioned above are complied with.[26] Consequently, no legal basis separate from that which allowed the original processing of the personal data is required.

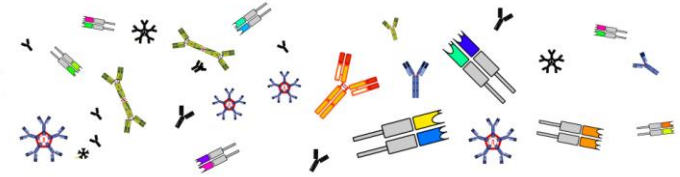
In addition, the controller remains bound by the GDPR's notice requirements. This notice should be updated when the controller intends to further process data for a different purpose, including research.[27] The notice requirement does however not apply when the provision of information proves to be impossible or would involve a disproportionate This particularly can be the case where processing is carried out for scientific research purposes, subject to appropriate safeguards. Another exemption to the notice requirement applies when the notice would be likely to render impossible or seriously impair the achievement of the research objectives, provided again that appropriate safeguards are in place, including making the information publicly available.[28]

Furthermore, in case of processing for research purposes, certain rights of the data subjects may also be carved out by the GDPR or Member State law.[29]

[22] The Article 29 Working Party was an advisory body made up of a representative from the data protection authority of each EU Member State, the European Data Protection Supervisor and the European Commission. On 25 May 2018, it has been replaced by the European Data Protection Board under the GDPR.
[23] WP29 Opinion 06/2014 on the notion of legitimate interests of the data controller under article 7 of Directive 95/46/EC, 24-25. Accessible at: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp217_en.pdf.
[24] Recital 47 GDPR.
[25] Art. 5(1) (b) GDPR.
[26] Recital 50 GDPR.
[27] Article 13 (3) and 14(4) GDPR.
[28] Recital 62 GDPR and article 14 (5) (b) GDPR.
[29] Article 17 (3) (d), 21 (6) and 89 (2) 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)



It is thus clear that personal data can be processed and further processed for scientific research purposes, subject to appropriate conditions and safeguards set out in Union or Member State law. Scientific research should in this context be interpreted broadly, including technological development and demonstration, fundamental research, applied research and privately funded research.[30]

2.3.2 Resulting requirements:

- iReceptor Plus should make sure that appropriate safeguards are put in place for the rights and freedoms of patients and research participants. These safeguards should ensure that technical and organisational measures are in place in particular in order to ensure respect for the principle of data minimization (e.g. pseudonymization, implementing an Information Security Management System, cryptographic protection during storage and transfer, authentication, authorization, physical and logical access to data, access logging…).

- iReceptor Plus should furthermore review national data protection laws to assess if more stringent requirements or conditions with regard to the further processing of genetic and health data apply.

2.4 Continuous risk assessment

2.4.1 Meaning and relevance to iReceptor Plus:

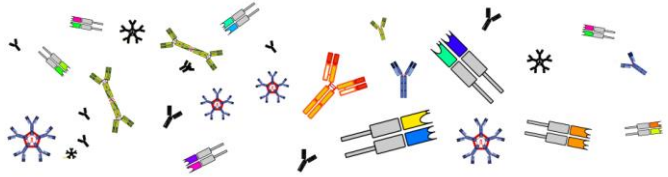
The activities performed in iReceptor Plus present elevated privacy, security and ethics risks. First of all, any research concerning the processing of genetic data (in non-anonymous form) presents significant risks, given that genetic data can be predictive of genetic predispositions concerning individuals and their family and may thus create threats of discrimination and stigmatization. Moreover, iReceptor Plus aims to apply a *big data analytics approach*, in which data sets from multiple repositories are combined. This big data analytics approach and the implementation of artificial intelligence (AI) techniques raises additional ethics questions.

More importantly, the research methodologies in iReceptor Plus are dynamic and developmental, in a way that anticipating all risks relating to the research endeavour at this stage is very difficult. For these reasons, in iReceptor Plus, risk should be monitored and evaluated on a continuous basis and risk assessment should respond to changes in the circumstances of the research setting.

This is also a principle of EU data protection law: the GDPR requires that, where a type of

[30] Recital 159 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)



processing in particular using new technologies, and taking into account the nature, scope, context and purposes of the processing, is likely to result in a high risk to the rights and freedoms of natural persons, the controller[31] shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations of personal data.[32] Such privacy impact assessments are also commonplace in many other jurisdictions.[33]

2.4.2 Resulting requirements:

- To the extent that personal data are processed in iReceptor Plus, a data protection impact assessment (DPIA) should be conducted and maintained in the course of iReceptor Plus, to identify risks and propose mitigation strategies.

- This DPIA should comply with the requirements of the GDPR but should also encompass any requirements resulting from other applicable laws and any other identified risks that go beyond data protection/privacy risks. Notably, the risks and potential mitigation strategies related to the big data analytics and deep learning approach in iReceptor Plus should be covered[34].

2.5 Privacy / data protection by design and default

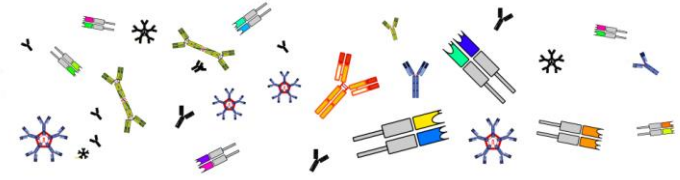
2.5.1 Meaning and relevance to iReceptor Plus:

Data protection by design and default are two key principles in addressing the ethics concerns that arise in the design phase of a project. In essence, these principles require that data protection is integrated into the processing and business activities from the design stage right through the lifecycle, and that, by default, only data that are necessary to achieve the specific purposes envisaged are processed.[35]

Data protection by design requires controllers[36], both at the time of the determination of the means of processing and at the time of the processing itself, to implement appropriate technical and organisational measures, such as pseudonymisation, which are designed to implement data protection principles, such as data minimisation, in an effective manner and to integrate the

[31] Under the GDPR, the controller is the individual or entity, that alone or jointly with others determines the purposes and means of the processing of personal data.
[32] Article 35 (1) GDPR.
[33] See for example: https://www.priv.gc.ca/en/privacy-topics/privacy-impact-assessments/, https://www.state.gov/privacy-impact-assessments-privacy-office/ .
[34] See section 2.8.
[35] https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/accountability-and-governance/data-protection-by-design-and-default/ .
[36] For a definition of controller, see footnote 9.

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)



necessary safeguards into the processing in order to meet the requirements of the GDPR and to protect the rights of data subjects.[37]

Data protection by default requires controllers to implement appropriate technical and organisational measures for ensuring that, by default, only personal data which are necessary for each specific purpose of the processing are processed.[38]

Although data protection by design and default are principles created and defined by the EU GDPR, at their origin, these concepts build on the principle of 'privacy by design' as introduced by the Information and Privacy Commissioner of Ontario, Dr. Ann Cavoukian in 1995.

The foundational principles of privacy by design[39] are:

- Proactive not reactive, preventative not remedial
- Privacy as default setting
- Privacy embedded into design
- Full functionality – positive sum, not zero-sum
- End-to-end security – full lifecycle protection
- Visibility and transparency
- Respect for user privacy

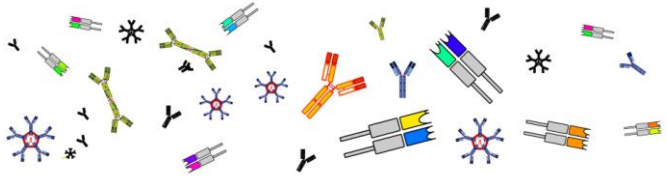
These principles were recognised by the International Conference of Data Protection and Privacy Commissioners (ICDPPC) in a resolution on privacy by design of 2010.[40] As such the principle of privacy by design is also accepted by plenty of other data protection authorities in different countries, such as the Federal Trade Commission of the US and the privacy commissioners of Ontario, British Columbia, Norway and Israel.

iReceptor Plus is aimed at developing a platform to integrate distributed repositories of AIRR-seq data. This may require complex, sensitive and large-scale data processing. It is therefore critical that measures be taken to mitigate the risks raised by the data processing activities in iReceptor Plus by default, rather than as an optional feature; and to enhance the level of data protection through the design of the platform, by respecting general, commonly accepted principles of data protection as well as the requirements of the applicable data protection laws from the outset of the project.

For example, the iReceptor Plus project software will include one or more web-based user interfaces (e.g. the iReceptor Scientific Gateway). These user interfaces will query the network of AIRR-seq repositories (using whatever authentication and authorization mechanisms required

[37] Article 25(1) GDPR.
[38] Article 25 (2) GDPR.
[39] Accessible at: https://iab.org/wp-content/IAB-uploads/2011/03/fred_carter.pdf.
[40] Accessible at: https://icdppc.org/wp-content/uploads/2015/02/32-Conference-Israel-resolution-on-Privacy-by-Design.pdf.

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)



by the repositories) to access, federate, and analyse data. Access to data from the repositories will not be possible without adhering to the repository's authentication and authorization mechanisms. Moreover, data will be protected while in transit between systems and while within iReceptor Plus systems using authentication, authorization, security, and data encryption best practices. The techniques used to protect data within the platform will be documented so that users who are accessing data with ethics constraints (e.g. users who have a data sharing agreement to access data from a specific repository) can make informed decisions as to whether the iReceptor Plus platform is suitable for their use. As with the iReceptor Plus repository software, the goal of the iReceptor Plus user interface and analysis platform in this context is to provide a privacy-friendly tool for analysing AIRR-seq data and to make it easy for both the users that are analysing the data and the data stewards that manage the data to make informed decisions as to whether the iReceptor Plus analysis platform is appropriate for performing their analyses in a secure and ethical manner.

2.5.2 Resulting requirements:

- The iReceptor Plus platform should be designed keeping in mind generally accepted principles of privacy / data protection by design and default and should implement appropriate measures to this end. This could include:
    - The anonymisation or pseudonymisation of personal data
    - Respect for the principles of purpose limitation and data minimisation
    - Ensuring accountability for all processing activities
    - Providing appropriate security of data (e.g. applied cryptography)
    - Providing access rights to individuals regarding their personal data
    - Etc.

  This requirement is at the heart of iReceptor Plus, which in essence has the goal of providing a tool for researchers to access and share AIRR-seq data while remaining compliant with all legal, privacy, ethics and security requirements.

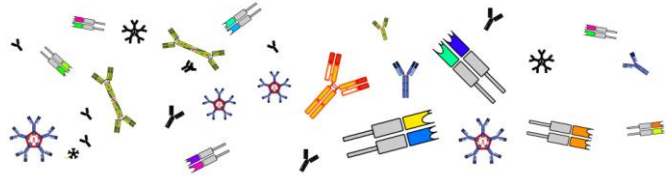
2.6 Transfer of personal data

2.6.1 Meaning and relevance to iReceptor Plus:

One of the fundamental characteristics of iReceptor Plus as a project, is that it will link distributed repositories of AIRR-seq and other data from all over the world into one platform. This distributed, federated approach implies that large amounts of (possibly) personal data will be shared cross-borders.

Cross-border sharing of personal data implies that data are transferred from one country to another. Countries with comprehensive data protection laws therefore may establish specific requirements in order for a transfer of personal data to a third country (or international

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)



organisation) to be allowed. These requirements relate to the idea that personal data transferred to a third country should be granted the same level of protection abroad, as it is awarded in the country from which the data are transferred. As such, the data protection offered to a natural person with regard to its personal data will travel with the data when it leaves the country.

In the EU, the GDPR, includes an entire chapter on the transfer of personal data to a third country or international organisation. 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

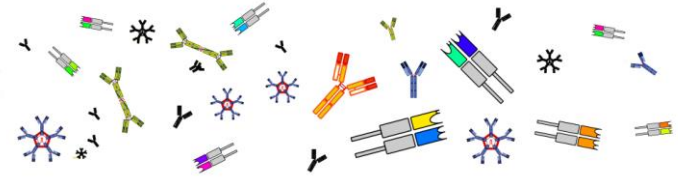
the international organisation in question ensures an *adequate level of protection*, i.e. if the EC has made an **adequacy finding** for the country or territory concerned. In that case, the transfer of personal data to that third country or international organisation will not require any specific authorisation.

➢ 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.

These appropriate safeguards may be provided for by:

- A legally binding and enforceable instrument between public authorities or bodies
- Binding corporate rules (BCR's)
- Standard data protection clauses (adopted by the Commission or adopted by a supervisory authority and approved by the Commission)
- An approved code of conduct (together with commitments of the data importer in the third country to apply the appropriate safeguards)
- An approved certification mechanism (together with commitments of the data importer in the third country to apply the appropriate safeguards)
- Contractual clauses between the data exporter and the data importer in the third country or international organisation, authorised by the competent supervisory authority.
- 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)



➢ 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:

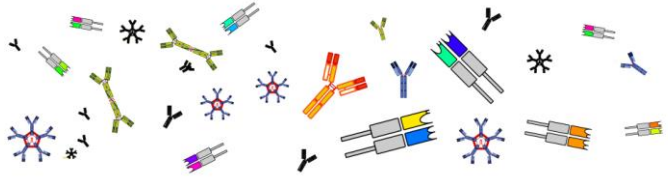
- the data subject has explicitly *consented* to the proposed transfer, after having been informed of the possible risks thereof in absence of adequacy decision or appropriate safeguards;
- the transfer is *necessary for the performance of a contract between the data subject and the controller* or the 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 last option will however not apply in iReceptor Plus, given that repetitive data transfers are envisaged.

Furthermore, it should not be forgotten that the transfer of personal data to a third country or international organisation in itself constitutes a processing activity, which requires a legal ground as well as fulfilling the conditions related to processing special categories of data, namely genetic and health data.

In other countries (not part of the EU), the requirements for the transfer of personal data to third countries are often either inspired by the GDPR, or mainly put emphasis on the principles of transparency, purpose limitation and consent.

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)



2.6.2 *Resulting requirements:*

- To the extent that personal data are processed in iReceptor Plus, it should be ensured that all national requirements relating to a transfer of personal data to third countries or international organisations are complied with. These requirements may differ among national laws.

2.7 Data security, retention and deletion

2.7.1 *Meaning and relevance to iReceptor Plus:*

Data security

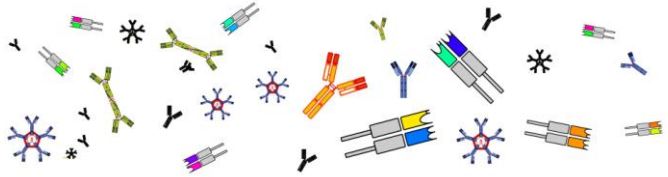
Data security refers to the layer of security in an information system that is devoted to adding protections to the data in the system itself and controlling the access to the data through identity and access management.

Because of the distributed, federated approach of the iReceptor Plus platform, addressing data security in iReceptor Plus is quite challenging. Indeed, the service-oriented architecture of the platform, which implies that the platform is composed of services from different sources (i.e. different repositories) entails that data security needs to be ensured at two levels: (1) the level concerning the security of the service itself, composed by its parts and (2) the level concerning the security of the contents of the systems, the data. Adding to this complexity is the particular sensitivity of the data processed in iReceptor Plus and the application of a big data analytics approach, which both require enhanced levels of security. Therefore, a holistic approach to data security is required, in which the data flows throughout all of the applications in the iReceptor Plus system are analysed and secured.

To this end, it is crucial that a system of identity and access management be set up to control the access to the data available through the iReceptor Plus platform. Through identification, authentication and authorization, the access to the data can be controlled in a way that only authorized persons access data with a certain level of confidentiality. This can be supplemented with security auditing and monitoring to detect suspicious behaviour or unauthorised changes.

Furthermore, the data itself should be protected to ensure confidentiality, integrity and non-repudiation. This can be done in multiple layers, for example by providing both storage encryption and data encryption. This not only protects the data against malicious attacks but may also provide a means for restricting the authorisation of users to a certain extent.

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)



By implementing such measures, iReceptor Plus can guarantee an appropriate security protection relative to the sensitivity of the information at stake, as is required under most data protection laws.[41]

Retention and deletion

If personal data are being processed, the retention and deletion of the data is just as important as the security of the data. One of the fundamental principles of most data protection laws states that personal data may only be kept for as long as is necessary for the purposes for which the data were collected (principle of storage limitation). However, according to EU law, personal data may be stored for longer periods insofar as the personal data will be processed solely for scientific research purposes and appropriate technical and organisational measures are taken to safeguard the rights and freedoms of the person concerned, such as anonymisation or pseudonymisation.

Given that the distributed repositories in iReceptor Plus remain in charge of the data curated in their repositories, these individual repositories are responsible for putting in place appropriate data retention and deletion policies and informing patients and research participants about these policies. From a technical perspective however, iReceptor Plus will have to ensure that no personal data is stored or can be accessed after the retention and deletion period of the individual repository curating the data has expired. If personal data with a certain level of confidentiality is however shared between repositories via a reciprocal data transfer agreement, then this agreement will have to cover the retention and deletion of these data by the repository receiving the data.

2.7.2 Resulting requirements:

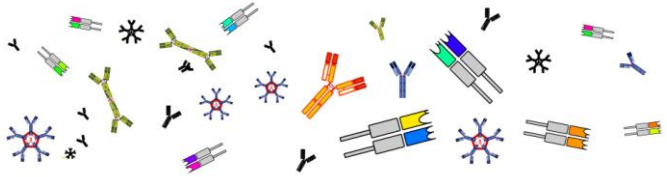
- iReceptor Plus needs to implement security requirements that are appropriate to the sensitivity of the information at and that take into account the system and its components as well as the data accessible through the system. This will among others require the encryption of data, an identity and access management policy (with authentication, authorisation and logging processes), an acceptable use policy and a data breach policy.

- All data processing activities undertaken through the iReceptor Plus platform must furthermore be subject to monitoring at the local level by the individual repositories, who must provide for appropriate data retention and deletion policies.

- iReceptor Plus must ensure that no personal data are available beyond the data retention period of the individual repositories and that any stored personal data is deleted following the data deletion policy of the individual repositories (to the extent that data is stored).

[41] See for example article 32 of the GDPR and principle 7 of the PIPEDA Fair Information Principles.

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)



2.8 Big data and artificial intelligence

2.8.1 Meaning and relevance to iReceptor Plus:

Big data refers to the field dealing with the processing, collection, storage and analysis of large amounts of structured and unstructured data, which is complex to process using traditional database and software techniques. Big data are characterized by their volume (scale of data), variety (produced by different sources in different formats) and velocity (connected to the analysis of streaming data).

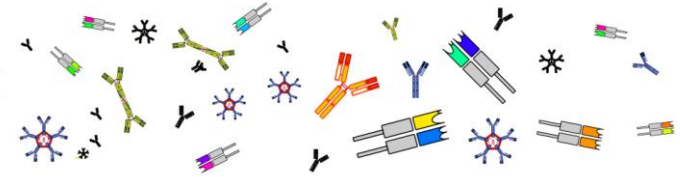
Artificial intelligence (AI) refers to systems that show intelligent behaviour. By analysing their environment these systems can perform various tasks with some degree of autonomy to achieve specific goals. In essence, artificial intelligence systems act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information derived from this data and deciding the best action(s) to take to achieve the specific goal that was given to the system.

Big data analytics relates to three concepts: big data, artificial intelligence and machine learning. Basically, big data is an asset that is difficult to exploit. Artificial intelligence (AI) is the key to unlocking the value of big data and machine learning is one of the technical mechanisms that underpins and facilitates this AI.

Big data analytics present great opportunities and can bring significant benefits to society, for example by offering new methods and solutions in various fields such as public health and medical care. However, big data analytics, notably machine learning and deep learning technologies, also pose challenges in regards the respect of the rights to privacy and data protection, ethical and human rights considerations. These challenges relate (amongst others) to biases resulting in discrimination and interference with individuals' rights (such as freedom of expression and information) or the exclusion of people from certain aspects of personal, social and professional life. In addition, big data analytics may run counter to the principles of fairness and transparency, purpose limitation and data minimisation.

In iReceptor Plus, the analysis of AIRR-seq data will be performed through artificial intelligence Natural Language Processing (NLP) and Deep Learning (DL). It is therefore important that iReceptor Plus, when designing, developing and using this artificial intelligence, fully respects human rights, in particular the right to the protection of personal data and privacy, as well as human dignity, non-discrimination and other fundamental values; and that it ensures that

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)



individuals maintain control and understanding of the artificial intelligence systems.

Therefore, it is advisable that iReceptor Plus follows the guidance[42] for trustworthy AI presented by the High-Level Expert Group on Artificial Intelligence, which puts forward 7 requirements for trustworthy AI:

- Human agency and oversight

  AI systems should support human autonomy and decision-making, as prescribed by the principle of respect for human autonomy. AI systems should support the user's agency, foster fundamental rights and allow for human oversight.

- Technical robustness and safety

  AI systems should be developed with a preventive approach to risks and minimising or preventing harm. They should be secure and resilient to attacks, should have safeguards that enable a fallback plan in case of problems and should produce accurate, reproducible and reliable results.

- Privacy and data governance

  AI systems should guarantee privacy and data protection throughout a system's entire lifecycle. Adequate privacy protection also necessitates data governance that covers the quality and integrity of the data used as well as data access protocols.

- Transparency

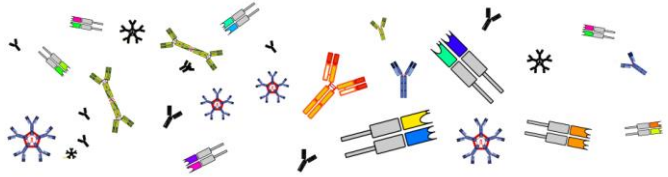
  AI systems should be transparent, in a way that data sets and processes are traceable and explainable. Moreover, people need to be informed when they are interacting with an AI system.

- Diversity, non-discrimination and fairness

  AI systems should ensure inclusion and diversity throughout the entire system's life cycle. It should avoid unfair biases which could lead to discrimination and should be designed in an accessible and universal way.

[42] Available for download at: https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai.

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)



- Environmental and societal well-being

  AI systems should be sustainable and ecologically responsible to the greatest extent possible. They should be used to benefit all human beings, including future generations.

- Accountability

  AI systems should be accountable. This necessitates that mechanisms be put in place to ensure responsibility and accountability for AI systems and their outcomes, both before and after their development, deployment and use.

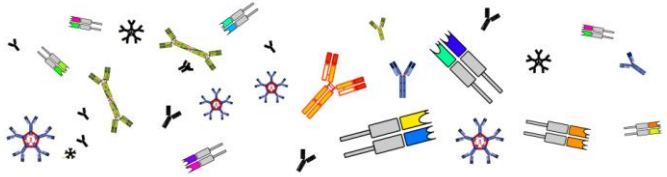
These requirements should be continuously evaluated and addressed throughout the AI system's lifecycle.

2.8.2 Resulting requirements:

- When designing, developing and using artificial intelligence, iReceptor Plus should assess the impact of this artificial intelligence on human rights, particularly the right to the protection of personal data, as well as human dignity and other fundamental values. This assessment should be part of a broader data protection impact assessment and should take into account the requirements set by the High-Level Expert Group on Artificial Intelligence. To this end, the 'Trustworthy AI Assessment List' incorporated in the guidance document of this group can be used.

- A project-internal ethics committee should be set up to assess the ethical aspects of implementing artificial intelligence and to provide general input for the data protection impact assessment.

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)



2.9 Data ownership and IP rights

2.9.1 Meaning and relevance to iReceptor Plus:

The growing importance of data, including personal information, has given rise to increased discussions about property rights in data.[43] These discussions relate (amongst others) to the question about how to strike a balance between the interests of the industry to protect their investment in both the collection and creation/generation of data, the interests of individuals to control (the collection of, access to and use of) their personal data and the interest of the public to access and re-use data.

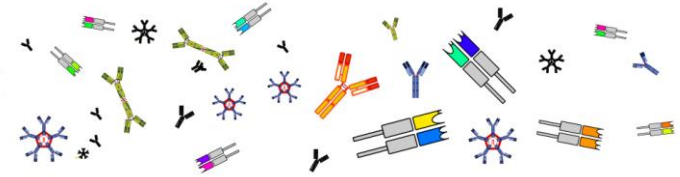
From a legal point of view, up until now, there does not exist a specific 'data ownership right'. Data ownership in this context should be understood to imply rights over a property such as being able to enjoy, use, sell, rent, give away or even destroy an item of property. It does not refer to the responsibility and accountability for specific databases that is assigned to a certain entity, so as to ensure data quality, curation, data protection and security throughout the life cycle of the data.[44] This last interpretation of data ownership, as often used by businesses should better be referred to as 'data stewardship' or 'data custodianship'.

Nevertheless, the commercial value and economic importance of data has led to calls for an ownership right in data. Creating such a sui generis data ownership right might however be challenging, given that (1) it is difficult to define what 'data' is , (2) it is difficult to 'locate' the ownership: different stakeholders may try to claim ownership in (parts of) datasets because they, for instance, create or generate data, or because they use, compile, select, structure, re-format, enrich or add value to the data, (3) it is difficult to establish significant rights of access and use, both in regards of the public and (possibly) the person to whom the data relate.[45]

As matters stand, data are protected in the sense that data that are kept confidential can be protected as confidential information. Moreover, compilations of data can be protected under copyright law if this compilation meets the threshold of originality. This protection however only applies to the original selection or arrangement, not to the underlying data. Lastly, in the European Union, a sui generis database right exists that offers more robust protection for compilations of data. Again, however, the protection does not extend to the data that make up such a compilation. Apart from these protection mechanisms, people who want to establish a form of ownership in data themselves are thus obliged to rely on contractual arrangements.

[43] T. SCASSA, Data Ownership, Centre for International Governance Innovation (CIGI) Papers No. 187.
[44] OECD, Data-driven innovation: Big Data for Growth and Well-being (OECD Publishing 2015), 195.
[45] T. SCASSA, Data Ownership, Centre for International Governance Innovation (CIGI) Papers No. 187.

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)



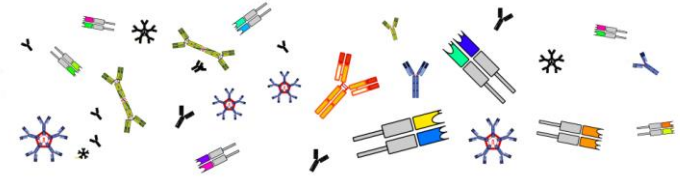
In iReceptor Plus specifically, the question arises who owns (1) the data in the distributed repositories and (2) the 'analysis data' that are generated by combining different datasets through the iReceptor Plus platform and applying artificial intelligence techniques. Moreover, several questions of intellectual property rights may also come into play in the iReceptor Plus project. First of all, in the 'creation phase', copyright protection as well as software protection (in the EU) may apply to the iReceptor Plus platform software and the iReceptor Plus Turnkey repository solution. Moreover, a trademark protection could be awarded to the iReceptor Plus project name and logo. In the 'data collection and presentation phase' of the project, however, there will most likely not be any applicable IP rights, given that data are collected by the individual repositories and there is no human creation in the data presented after analysis by artificial intelligence techniques. Lastly, in the 'access phase' of the project, the question could be asked if any downstream intellectual property rights should apply to the inventions and works created on the basis of the data extracted from the iReceptor Plus platform.

All of these questions should be meticulously addressed in an intellectual property rights management plan, which takes into account the provisions of the grant agreement and consortium agreement as well as the background of the iReceptor Plus partners. Where necessary, contractual agreements should be set up with the individual repository owners and IP license terms should be imposed upon the platform users.

2.9.2 Resulting requirements:

- iReceptor Plus should draw up a detailed intellectual property rights management plan to address all intellectual property rights questions arising throughout the project and should formulate contractual arrangements and terms regarding data ownership and intellectual property rights. These documents should assess the need/strategy for protection of each iReceptor Plus output and should in any case comply with the provisions of the grant and consortium agreement and should respect the existing background of the partners.

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)



2.10 National compliance

2.10.1 Meaning and relevance to iReceptor Plus:

As matters stand, the iReceptor Plus project will include repositories based in the EU (France, Germany), Israel, Norway, Canada (British Columbia and Ontario) and the United States of America (California and Texas). It is however envisaged that other repositories, from non-beneficiaries in this project will also be participating in the future.

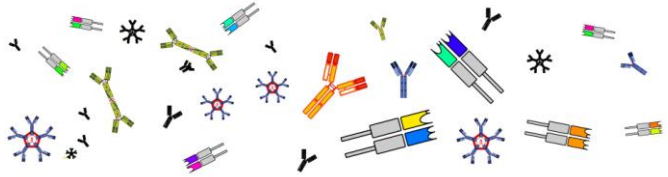
Although EU laws as well as Canadian and American federal laws include harmonising rules, privacy and data protection law, ethics and security requirements in the health sector remain to a large extent a matter of national or state law and interpretation. This is the case for certain key topics such as the requirements for the processing of data concerning health, the definition of appropriate safeguards for scientific research, consent requirements, the need for national certifications… All of these topics are critical for ensuring the legal compliance of iReceptor Plus, and as such a thorough analysis of national and state-specific requirements is necessary.

2.10.2 Resulting requirements:

- The sharing of data by individual repositories with the iReceptor Plus network must be subject to national compliance supervision and approval from an ethics committee which is familiar with national compliance requirements.

- National and state compliance requirements for iReceptor Plus should be further identified and detailed throughout the duration of the project, so as to ensure strict compliance.

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 Conclusion

This first version of the 'Legal, ethics, privacy & security framework' (D8.2) defines the initial legal, ethical, privacy and security requirements for iReceptor Plus at a high level. It is aimed at reviewing and structuring key challenges and developing a relevant project policy towards resolving them, across 10 specific domains:

1. Anonymisation and pseudonymisation
2. Informed consent, fairness and transparency
3. Lawfulness and further processing of personal data
4. Continuous risk assessment
5. Data protection by design and default
6. Transfer of personal data
7. Data security, deletion and archiving
8. Big data and artificial intelligence
9. Data ownership and IP rights
10. National compliance

In each of these domains, the high-level issue is reflected upon in light of iReceptor Plus and is then linked to several requirements for the project. In this manner, the deliverable can act as an elementary legal and ethical guide for the activities performed in iReceptor Plus.

As noted in the introduction, this legal, ethics, privacy & security framework should not be seen as a fully mature and exhaustive document, but rather as a first outline of key challenges deemed relevant for further project work. The legal, ethical, privacy and security requirements set forth in this deliverable will be further addressed in future iterations (M24 and M36) as well as in other deliverables, such as D3.3 (Monitoring and GDPR aspects) and D3.4 (Recommendations and proposals for new regulatory regimes), based on additional input from iReceptor Plus partners and the further identification of relevant factors and regulatory processes at the EU level and in each of the countries participating.

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)