Project Acronym: Nutrishield
Grant Agreement number: 818110 (H2020-SFS-2018-IA)
Project Full Title: Fact-based personalised nutrition for the young

DELIVERABLE

D6.5 – Data governance and ETL specifications -Preliminary Version

Dissemination level	PU - Public
Type of Document	Report
Contractual date of delivery	29/02/2020
Deliverable Leader	Vrije Universiteit Brussels (VUB)
Status & version	Final, v1.0 – 29/02/2020
WP responsible	WP6 (INTRA)
Keywords:	ETL specifications, personal data protection

Deliverable Leader:	VUB
Contributors:	Dimitra Markopoulou, Evangelos Papakonstantinou (VUB)
Reviewers:	Ioannis Sarris (INTRA), Ioannis Daskalopoulos (INTRA), Spyros Evangelatos (INTRA), Eirini Bathrellou (HUA)
Approved by:	Miltos Vasiliadis (ALPES)

Document History			
Version	**Date**	**Contributor(s)**	**Description**
v0.1	01/11/2019	VUB	Draft
v0.5	01/02/2020	VUB	First complete version
v0.6	26/02/2020	INTRA	Reviewers' comments
v0.7	26/02/2020	HUA	Reviewers' comments
v0.9	28/02/2020	VUB	Reviewers' comments incorporated
v1.0	29/02/2020	ALPES	Final version

This document is part of a project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 818110.

Executive Summary

The focus of the present report is to describe and evaluate the ETL procedures that will be employed and established by the project partners, in the context of the NUTRISHIELD project. It is therefore examined whether this procedure is compliant with the main personal data processing principles indicated by the GDPR. Given the project's specifications and description as a health (with emphasis on nutrition) tool, special attention is being attributed to the protection of personal data, in particular, health and genetic data, that may potentially be processed during the project's duration. Having this in mind, the analysis that follows will endeavor to complement the project partners' effort to design the NUTRISHIELD dashboard and mobile application while abiding to the principles of data protection by design and by default and to provide them with useful guidelines throughout the project's execution, in particular as regards the protection of the rights of the patients/participants in the NUTRISHIELD research (clinical studies). The present report represents a preliminary analysis on any issues relating to the data processing activities that are anticipated to take place for the purposes of NUTRISHIELD. As the project progresses and the processing activities are finalized, an updated version of this report (due M24) will address any outstanding issues.

Table of Contents

1. Personal data protection under the EU Regulatory framework – Regulation 2016/679 (GDPR)

1.1. The EU regulatory framework

The General Data Protection Regulation[1] was adopted in May 2016 in replacement of Directive 95/46/EC (the Data Protection Directive)[2], that was, until then, the main tool for protecting natural persons against unlawful processing of their personal data. The Regulation became fully enforceable in the European Union in May 2018. Contrary to the Directive, the GDPR is a regulatory tool of catholic and direct effect that intends to address any inconsistency in national laws and to succeed a harmonised personal data protection approach among Member States.
GDPR regulates the processing by an individual, a company or an organization of personal data relating to individuals in the EU. The Regulation does not apply to the processing of personal data of deceased persons or of legal entities. Its provisions do not apply to data processing by an individual for purely personal reasons or for activities carried out in one's home provided there is no connection to a professional or commercial activity.
At the same time the protection of natural persons in relation to the processing of personal data is a fundamental right. Article 8(1) of the Charter of Fundamental Rights of the European Union (the 'Charter') and Article 16(1) of the Treaty on the Functioning of the European Union (TFEU) provide that everyone has the right to the protection of personal data concerning him or her.

1.2. Definitions of basic concepts under the GDPR

a) Personal Data

The definition of "personal data" is included in Article 4(1) of the GDPR: **personal data** means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as

[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 personal data and on the free movement of such data, and repealing Directive 95/46/EC

[2] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data (replaced by the GDPR)

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.

The notion of identifiability is further addressed under recital 26 of the Regulation 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*". The recital clarifies that personal data which have undergone pseudonymisation, which could be attributed to a natural person by the use of additional information, should be considered to be information on an identifiable natural person.

The GDPR does not apply to anonymous information, that is information which does not relate to an identified or identifiable natural person. The same rule governs any data that were rendered anonymous in such a manner that the data subject is not or no longer identifiable.

b) "Processing" of personal data

A definition of "**processing**" of personal data is provided under Article 4(2) of the Regulation. 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*".

c. Pseudonymisation

Pseudonymisation means the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person.

d. Controllers – processors – joint controllers – recipients

The definition of a **controller** is provided under article 4(7) of the GDPR. According to said provision controller means "*the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data; where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law*". New addition introduced by the Regulation is the explicit reference to the notion of joint controllers. Article 26 of the Regulation states that "*Where two or more controllers jointly determine the purposes and means of processing, they shall be joint controllers. They shall in a transparent manner determine their respective responsibilities for compliance with the obligations under this Regulation* […]".

Article 4(8) defines a **processor** as "a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller".

Finally, a **recipient** is defined under article 4(9). In particular, "recipient' means a natural or legal person, public authority, agency or another body, to which the personal data are disclosed, whether a third party or not".

e) Special categories of personal data and their processing

Special attention should be given to categories of data that do not fall under the generic definition of personal data mentioned above but to a specific group that of special categories of data. It is noted that the term sensitive data that was used in the Directive, is replaced in the new Regulation by the term "**special categories of personal data**". These include:

- data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership
- genetic data that include personal data relating to the inherited or acquired genetic characteristics of a natural person which give unique information about the physiology or the health of that natural person and which result, in particular, from an analysis of a biological sample from the natural person in question**;**
- biometric data that include personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, which allow or confirm the unique identification of that natural person, such as facial images or dactyloscopic data; and
- data concerning health that refer to personal data related to the physical or mental health of a natural person, including the provision of health care services, which reveal information about his or her health status. The first two categories constitute additions in the data protection field that come as a result of scientific developments in their respective fields.

Article 9 (1) of the GDPR introduces a general prohibition as regards this category of data. The Article reads as follows: "processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation shall be prohibited".

There are exceptions though to this general rule that are included in par. 2 of the same Article 9, and are outlined below:

a. The data subject has given explicit consent to the processing of those personal data for one or more specified purposes, except where Union or Member State law provide that the prohibition referred to in paragraph 1 may not be lifted by the data subject;

b. Processing is necessary for the purposes of carrying out the obligations and exercising specific rights of the controller or of the data subject in the field of employment and social security and social protection law in so far as it is authorised by Union or Member State law or a collective agreement pursuant to Member State law providing for appropriate safeguards for the fundamental rights and the interests of the data subject;
c. Processing is necessary to protect the vital interests of the data subject or of another natural person where the data subject is physically or legally incapable of giving consent;
d. Processing is carried out in the course of its legitimate activities with appropriate safeguards by a foundation, association or any other not-for-profit body with a political, philosophical, religious or trade union aim and on condition that the processing relates solely to the members or to former members of the body or to persons who have regular contact with it in connection with its purposes and that the personal data are not disclosed outside that body without the consent of the data subjects;
e. Processing relates to personal data which are manifestly made public by the data subject;
f. processing is necessary for the establishment, exercise or defence of legal claims or whenever courts are acting in their judicial capacity;
g. Processing is necessary for reasons of substantial public interest, on the basis of Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject;
h. Processing is necessary for the purposes of preventive or occupational medicine, for the assessment of the working capacity of the employee, medical diagnosis, the provision of health or social care or treatment or the management of health or social care systems and services on the basis of Union or Member State law or pursuant to contract with a health professional and subject to the conditions and safeguards referred to in paragraph 3;
i. Processing is necessary for reasons of public interest in the area of public health, such as protecting against serious cross-border threats to health or ensuring high standards of quality and safety of health care and of medicinal products or medical devices, on the basis of Union or Member State law which provides for suitable and specific measures to safeguard the rights and freedoms of the data subject, in particular professional secrecy;
j. Processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) based on Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject.

Given NUTRISHIELD's focus on processing of health data by medical professionals participating in the project, paragraph (h) of article 9 is thoroughly examined below under 2.3.

1.3. Principles related to personal data processing under the GDPR

Principles relating to processing of personal data are included in article 5 of the GDPR. The article reads as follows:

1. Personal data shall be:

a. processed lawfully, fairly and in a transparent manner in relation to the data subject ('lawfulness, fairness and transparency')

b. collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes; further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes ('purpose limitation');

c. adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ('data minimisation');

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

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

f. processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organizational measures ('integrity and confidentiality').

2. The controller shall be responsible for, and be able to demonstrate compliance with, paragraph 1 ('accountability').

To sum up, the processing principles provided under the General Data Protection Regulation are the principles of:

- **lawfulness, fairness and transparency**
- **purpose limitation**
- **data minimisation**
- **accuracy:**
- **storage limitation**
- **integrity and confidentiality**
- **accountability**

1.4. Lawfulness of processing in particular

Article 6 of the GDPR mentions the conditions that need to be observed in order for processing of personal data to be lawful. In short, these include:

- consent,
- performance of a contract,
- compliance with a legal obligation,
- protection of vital interests of the data subjects,
- public interest,
- overriding interest of the controller

These six legal grounds apply alternatively and not cumulatively. This does not exclude the possibility of two or more legal grounds to apply at the same time. The application of one of these six legal bases must be established prior to the processing activity and in relation to a specific purpose. However, if a controller chooses to rely for instance on consent for any part of the processing, it must be prepared to respect that choice and stop that part of the processing if an individual withdraws consent. Therefore, in this case the controller cannot swap from consent to other lawful bases.

1.5. What is considered valid consent under the GDPR

a. Definition of consent

When it comes to personal data processing, individual consent is arguably the most important legal ground for doing so lawfully.

The GDPR regulates consent in several articles. In particular:

A definition of consent is provided under article 4(11) of the GDPR: “consent of the data subject means any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her”.

The basic requirements for the effectiveness of a valid legal consent are defined in article 7 and further specified in recital 32 of the GDPR. In particular based on the definition of consent:

Consent must be freely given, specific, informed and unambiguous. In order to obtain freely given consent, it must be given on a voluntary basis.

- **Free** implies a real choice of the data subject

- For consent to be **informed** and **specific**, the data subject must at least be notified about the controller's identity, what kind of data will be processed, how it will be used and the purpose of the processing operations. Furthermore, the subject must be informed of his/her right to withdraw consent and how to exercise such right. The consent must be bound to one or several specified purposes which must then be sufficiently explained.
- The element of **unambiguous** in consent means that consent requires either a statement or a clear affirmative action. Consent cannot be implied.

Recital 32 of the GDPR further clarifies the specific criteria mentioned above in stating that: "*Consent should be given by a clear affirmative act establishing a freely given, specific, informed and unambiguous indication of the data subject's agreement to the processing of personal data relating to him or her, such as by a written statement, including by electronic means, or an oral statement. This could include ticking a box when visiting an internet website, choosing technical settings for information society services or another statement or conduct which clearly indicates in this context the data subject's acceptance of the proposed processing of his or her personal data. Silence, pre-ticked boxes or inactivity should not therefore constitute consent. Consent should cover all processing activities carried out for the same purpose or purposes. When the processing has multiple purposes, consent should be given for all of them. If the data subject's consent is to be given following a request by electronic means, the request must be clear, concise and not unnecessarily disruptive to the use of the service for which it is provided*".

Conditions for consent to be valid as a legal ground for lawful processing are listed in article 7 of the GDPR. In more detail:

- The controller shall be responsible to demonstrate that the data subject has consented to processing of his or her personal data;
- If consent is given in the context of a written declaration, which also concerns other matters, the request for consent shall be presented in a manner which is clearly distinguishable from the other matters;
- The data subject shall be free to withdraw his/her consent at any time;
- When the performance of a contract is conditional on consent to the processing of personal data that is not necessary for the performance of that contract, it should always be examined whether the consent has indeed been provided freely;

b. Consent for processing of special categories of personal data.

As regards special categories of personal data, article 9 of the GDPR specifically mentions that the data subject needs to provide his/her explicit consent to the processing of this category of personal data in order for the general prohibition of non-processing to not apply. The term explicit refers to the way consent is expressed by the data subject. It means that the data subject must give an express statement of consent. An obvious way to make sure consent is explicit would be to expressly confirm consent in a written statement.

According to the explanatory memorandum of the Regulation[3] "the criterion 'explicit' is added to avoid confusing parallelism with 'unambiguous' consent"

c. Children's consent

Children's consent, in particular, is regulated under article 8 of the Regulation. The article specifically deals with children's consent in relation to information society services. As regards the age limit for the provision to apply it states that the processing of the personal data of a child shall be lawful where the child is at least 16 years old. Where the child is below the age of 16 years, such processing shall be lawful only if and to the extent that consent is given or authorised by the holder of parental responsibility over the child. Recital 38 clarifies the above by stating that "*Children merit specific protection with regard to their personal data, as they may be less aware of the risks, consequences and safeguards concerned and their rights in relation to the processing of personal data. Such specific protection should, in particular, apply to the use of personal data of children for the purposes of marketing or creating personality or user profiles and the collection of personal data with regard to children when using services offered directly to a child*".

1.6. The principle of transparency

Although the term transparency is not defined in the GDPR, recital 39 provides some clarity explaining that "…*It should be transparent to natural persons that personal data concerning them are collected, used, consulted or otherwise processed and to what extent the personal data are or will be processed. The principle of transparency requires that any information and communication relating to the processing of those personal data be easily accessible and easy to understand, and that clear and plain language be used. That principle concerns, in particular, information to the data subjects on the identity of the controller and the purposes of the processing and further information to ensure fair and transparent processing in respect of the natural persons concerned and their right to obtain confirmation and communication of personal data concerning them which are being processed…*"

The article 29 Working party has issued a set of guidelines on transparency under the Regulation. According to these guidelines, transparency applies to three central areas:

- the provision of information to data subjects related to fair processing
- how data controllers communicate with data subjects in relation to their rights under the GDPR
- how data controllers facilitate the exercise by data subjects of their rights

[3] See proposal for a Regulation on the protection of individuals with regard to the processing of personal data and on the free movement of such data https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52012PC0011

According to the same guidelines the key articles in relation to transparency in the GDPR are found in chapter III (rights of the data subjects). In particular, article 12 sets out the general rules which apply to:

- the provision of information to data subjects (under Articles 13 - 14);
- communications with data subjects concerning the exercise of their rights (under Articles 15 - 22); and
- communications in relation to data breaches (Article 34).

These provisions are examined below under the relevant sections.

1.7. Processing of personal data in research

a. Scientific research under the GDPR

Prior to the GDPR Directive 95/46 recognised research as an important area of public interest justifying derogations from the general rules[4]. It left data protection in the areas of health and medical research largely to Member States to legislate nationally. The GDPR however deviated from this approach and introduced a more specific model for regulating personal data processing that takes place for research purposes. The term scientific research is not defined in the GDPR. It is referred to in recital 159 of the GDPR which states that, **where personal data are processed for scientific research purposes, this Regulation should also apply to that processing. Therefore, each of the principles under article 5 of the GDPR apply to all data processing including processing for research purposes**. The GDPR assumes a broad conception of research, including technological development, fundamental and applied research and privately funded research and 'studies conducted in the public interest in the area of public health.

The special regime in the GDPR for scientific research is handled specifically under Article 89. The article sets the safeguards and derogations relating to processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes and reads as follows:

[4] See for instance Recital 34 to Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data (replaced by the GDPR): "…*Member States must also be authorized, when justified by grounds of important public interest, to derogate from the prohibition on processing sensitive categories of data where important reasons of public interest so justify in areas such as public health and social protection - especially in order to ensure the quality and cost-effectiveness of the procedures used for settling claims for benefits and services in the health insurance system - scientific research and government statistics; whereas it is incumbent on them, however, to provide specific and suitable safeguards so as to protect the fundamental rights and the privacy of individuals*"

1. Processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, shall be subject to appropriate safeguards, in accordance with this Regulation, for the rights and freedoms of the data subject. Those safeguards shall ensure that technical and organisational measures are in place in particular in order to ensure respect for the principle of data minimisation. Those measures may include pseudonymisation provided that those purposes can be fulfilled in that manner. Where those purposes can be fulfilled by further processing which does not permit or no longer permits the identification of data subjects, those purposes shall be fulfilled in that manner.

2. Where personal data are processed for scientific or historical research purposes or statistical purposes, Union or Member State law may provide for derogations from the rights referred to in Articles 15, 16, 18 and 21 subject to the conditions and safeguards referred to in paragraph 1 of this Article in so far as such rights are likely to render impossible or seriously impair the achievement of the specific purposes, and such derogations are necessary for the fulfilment of those purposes.

3. Where personal data are processed for archiving purposes in the public interest, Union or Member State law may provide for derogations from the rights referred to in Articles 15, 16, 18, 19, 20 and 21 subject to the conditions and safeguards referred to in paragraph 1 of this Article in so far as such rights are likely to render impossible or seriously impair the achievement of the specific purposes, and such derogations are necessary for the fulfilment of those purposes.

4. Where processing referred to in paragraphs 2 and 3 serves at the same time another purpose, the derogations shall apply only to processing for the purposes referred to in those paragraphs".

Article 9(2)j of the GDPR introduces another exception when processing of special categories of data is conducted in the context of scientific research. The article in particular permits derogations to the prohibition of the processing of special categories of data when the purpose of scientific research is met. The article reads as follows: "*Paragraph 1 shall not apply if one of the following applies*: *processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) based on Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject".*

b. The EDPS opinion on application of the GDPR on scientific research

It is therefore evident that the GDPR assigns to scientific research a special regime, but otherwise leaves it to Member States to regulate it through national (GDPR-implementing) acts. Accordingly, there have been few guidelines or comprehensive studies on the application of data protection rules to research. On this basis, the EDPS published a preliminary opinion in an effort to highlight the challenges in the application of the GDPR to scientific research[5].

[5] See a preliminary opinion on data protection and scientific research by the EDPS, 6 January 2020 https://edps.europa.eu/sites/edp/files/publication/20-01-06_opinion_research_en.pdf

The opinion points out that "*All the provisions above outline a special regime for scientific research and demonstrate that research occupies a privileged position within the GDPR. This flexibility afforded to Member States through the provisions cited above, absent harmonised EU law except in a few areas (such as for clinical trials), means that the full extent of this special regime is not precisely delineated. Nevertheless, the special regime cannot be applied in such a way that the essence of the right to data protection is emptied out, and this includes data subject rights, appropriate organisational and technical measures against accidental or unlawful destruction, loss or alteration, and the supervision of an independent authority. Personal data which are 'publicly available' - such as those collected from social media sites - are still personal data. Any limitations to fundamental rights in law are to be interpreted restrictively and cannot be abused. It might be considered abusive for instance for a research organisation to interpret these special provisions in the GDPR as allowing the retention of personal data for indefinite periods and to deny data subjects rights to information. Further work is taking place on these questions within the EDPB and at national level".*

c. Consent in the context of personal data processing and consent of human participants in research

When consent is the legal basis for conducting research in accordance with the GDPR, this consent for the use of personal data should be distinguished from other consent requirements that serve as an ethical standard or procedural obligation.

The requirement of specific consent is somehow relaxed under recital 33 of the Regulation. Recital 33 states: "*It is often not possible to fully identify the purpose of personal data processing for scientific research purposes at the time of data collection. Therefore, data subjects should be allowed to give their consent to certain areas of scientific research when in keeping with recognised ethical standards for scientific research. Data subjects should have the opportunity to give their consent only to certain areas of research or parts of research projects to the extent allowed by the intended purpose."* This does not mean that scientific research must not have a well-described purpose, but that the purpose may described at a more generic level.

As mentioned by the Article 29 Working Party Guidelines on consent under Regulation 2016/679, [6] "*transparency is an additional safeguard when the circumstances of the research do not allow for a specific consent. A lack of purpose specification may be offset by information on the development of the purpose being provided regularly by controllers as the research project progresses so that, over time, the consent will be as specific as possible".*

In the recent opinion published by the EDPS[7], an effort is made to draw the line between consent as a legal basis for data protection – as this was analysed above- and consent of human participants as regards their participation in research. It is worth quoting the relevant paragraph: "*There is clear overlap between informed consent of human participants in research projects involving humans and consent under data protection law. But to view them as a single and indivisible requirement would be simplistic and misleading. Consent serves not only as a possible legal basis for the activity, it is also a safeguard - a means for giving*

[6] See Article 29 Working Party Guidelines on consent under Regulation 2016/679 adopted on 28 November 2017 https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=623051

[7] See footnote 2 above

individuals more control and choice and thereby for upholding society's trust in science. There may be circumstances in which consent is not the most suitable legal basis for data processing, and other lawful grounds under both Articles 6 and 9 GDPR should be considered. However, even where consent is not appropriate as a legal basis under GDPR, informed consent of a human research participant could still serve as an 'appropriate safeguard' of the rights of the data subject. Under what conditions such informed consent might be deemed an appropriate safeguard is still unclear. Certainly, innovative forms of consent in research activities, like tiered and dynamic consent, are promising practices that should be further encouraged and developed. The notion of consent in the two areas requires further discussion between the research community and data protection experts as part of a wider reflection on the role of consent and respect for individuals in the area of scientific research in the digital age".

d. Informed consent under Regulation on clinical trials

Informed consent is also regulated under Regulation No 536/2014[8] on clinical trials. Its article 29 states that:

"1. Informed consent shall be written, dated and signed by the person performing the interview referred to in point (c) of paragraph 2, and by the subject or, where the subject is not able to give informed consent, his or her legally designated representative after having been duly informed in accordance with paragraph 2. Where the subject is unable to write, consent may be given and recorded through appropriate alternative means in the presence of at least one impartial witness. In that case, the witness shall sign and date the informed consent document. The subject or, where the subject is not able to give informed consent, his or her legally designated representative shall be provided with a copy of the document (or the record) by which informed consent has been given. The informed consent shall be documented. Adequate time shall be given for the subject or his or her legally designated representative to consider his or her decision to participate in the clinical trial.

2. Information given to the subject or, where the subject is not able to give informed consent, his or her legally designated representative for the purposes of obtaining his or her informed consent shall:

a. enable the subject or his or her legally designated representative to understand:

i) the nature, objectives, benefits, implications, risks and inconveniences of the clinical trial

ii) the subject's rights and guarantees regarding his or her protection, in particular his or her right to refuse to participate and the right to withdraw from the clinical trial at any time without any resulting detriment and without having to provide any justification;

iii) the conditions under which the clinical trial is to be conducted, including the expected duration of the subject's participation in the clinical trial; and

iv) the possible treatment alternatives, including the follow-up measures if the participation of the subject in the clinical trial is discontinued;

b. be kept comprehensive, concise, clear, relevant and understandable to a layperson

c. be provided in a prior interview with a member of the investigating team who is appropriately qualified according to the law of the Member State concerned

d. include information about the applicable damage compensation system referred to in Article 76(1) and

[8] Regulation (EU) No 536/2014 of the European Parliament and of the Council of 16 April 2014 on clinical trials on medicinal products for human use, and repealing Directive 2001/20/EC Text with EEA relevance

e. include the EU trial number and information about the availability of the clinical trial results in accordance with paragraph 6.he information referred to in paragraph 2 shall be prepared in writing and be available to the subject or, where the subject is not able to give informed consent, his or her legally designated representative.

4. In the interview referred to in point (c) of paragraph 2, special attention shall be paid to the information needs of specific patient populations and of individual subjects, as well as to the methods used to give the information.

5. In the interview referred to in point (c) of paragraph 2, it shall be verified that the subject has understood the information

6. The subject shall be informed that the summary of the results of the clinical trial and a summary presented in terms understandable to a layperson will be made available in the EU database, referred to in Article 81 (the 'EU database'), pursuant to Article 37(4), irrespective of the outcome of the clinical trial, and, to the extent possible, when the summaries become available.

7. This Regulation is without prejudice to national law requiring that both the signature of the incapacitated person and the signature of his or her legally designated representative may be required on the informed consent form.

8. This Regulation is without prejudice to national law requiring that, in addition to the informed consent given by the legally designated representative, a minor who is capable of forming an opinion and assessing the information given to him or her, shall also assent in order to participate in a clinical trial.

Informed consent serves as the main safeguard for research participants' rights and freedoms. It is no wonder therefore why the conditions of informed consent are set thoroughly in different regulatory instruments. Acquiring a valid informed consent of the research participant and keeping this consent valid throughout the process should be the main concern of all parties involved in research.

1.8. Rights afforded to individuals (data subjects) under the GDPR

The rights of the data subjects are regulated in articles 13-21 of the Regulation and can be summarised as follows:

- The right to information
- The right to access the data
- The right to rectification

- The right to erasure (the right to be forgotten)
- The right to restriction of processing
- The right to data portability
- The right to object

a. The right to information

The right to information is regulated in two articles, namely Articles 13 and 14. Distinction is made between cases where the information was obtained from the data subject and other cases. In this context, article 13 regulates the case where personal data have been collected form the data subject. In this case, the controller shall at the time when personal data are obtained, provide the data subject with the following information:

a) the identity and the contact details of the controller and, where applicable, of the controller's representative;

b) the contact details of the data protection officer, where applicable;

c) the purposes of the processing for which the personal data are intended as well as the legal basis for the processing;

d) where the processing is based on point (f) of Article 6(1), the legitimate interests pursued by the controller or by a third party;

e) the recipients or categories of recipients of the personal data, if any;

f) where applicable, the fact that the controller intends to transfer personal data to a third country or international organisation and the existence or absence of an adequacy decision by the Commission, or in the case of transfers referred to in Article 46 or 47, or the second subparagraph of Article 49(1), reference to the appropriate or suitable safeguards and the means by which to obtain a copy of them or where they have been made available.

Paragraph 2 of Article 13 lists the additional information the controller needs to provide to the data subject when collecting his/her personal data, such as the period for which the personal data will be stored, the existence of the right to request access to the data or erasure, the right to withdraw consent at any time etc.

Article 14 lists the information to be provided to the data subject where personal data have not been obtained from the data subject itself. Paragraph 5 of article 14 sets some exemptions of the controllers' obligation to provide information, for instance when the provision of such information proves impossible or would involve a disproportionate effort or where personal data must remain confidential etc.

b. The right to access the data

The right of access by the data subject is regulated under article 15 of the Regulation. The data subject shall have the right to obtain from the controller confirmation as to whether or not personal data concerning him/her are being processed and if yes, access to such data as well as information regarding, among others, the purpose of processing, the recipients to whom the data have been or will be disclosed the existence of the right to request rectification, the right to lodge a complaint and others, the right to request rectification

etc. Paragraph 3 of article 15 sets the subject's right to request a copy of his/her personal data from the controller.

It is noted that the right to rectification is regulated separately in article 16. In particular, the data subject shall have the right to obtain from the controller without undue delay the rectification of inaccurate personal data concerning him or her.

c. The right to erasure (right to be forgotten)

Article 17 of the Regulation grants individuals the right to have their personal information deleted by data controllers, if specific conditions, as these are listed in its paragraph 1, are met. For instance, the personal data have been unlawfully processed or they are no longer necessary in relation to the purpose for which they were collected, or the data subject has withdrawn his/her consent and others. In the event that the controller has made such data public, reasonable steps (including technical measures) will be taken to notify controllers who are processing the personal data accordingly. Finally, the "right to be forgotten" (actually, to erasure of data) will not be applicable if it contrasts with the rights of freedom of expression and information as well as for several other legal grounds (compliance with a legal obligation, public interest, archiving purposes, etc., as set in paragraph 3).

d. The right to restriction of the processing

Article 18 of the Regulation regulates the right to restriction of the personal data processing. The conditions under which a data subject may exercise his/her rights are listed in the first paragraph of article 18 and include, for instance, the contest by the data subject of the accuracy of the personal data processed by the controller or the claim that the processing is unlawful and therefore the data subject opposes the erasure of his/her personal data. Recital 67 mentions some methods the controller may use to restrict the processing of personal data, such as, temporarily moving the selected data to another processing system, making the selected personal data unavailable to users, or temporarily removing published data from a website.

e. The right to data portability

Data portability is dealt with under article 20 of the GDPR and includes the data subject's right to receive the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and machine-readable format and have the right to transmit those data to another controller without hindrance from the controller to which the personal data have been provided. The right to data portability is provided to data subjects under two conditions:

a. the processing is carried out by automated means;

b. the processing is based on consent or on a contract.

f. The right to object

The right to object is laid down in Article 21 of the GDPR. Recital 69 of the Regulation clarifies the conditions under which a data subject may object to his/her data being processed. The recital reads as follows: "*Where personal data might lawfully be processed because 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, or on grounds of the legitimate interests of a controller or a third party, a data subject should, nevertheless, be entitled to object to the processing of any personal data relating to his or her particular situation. It should be for the controller to demonstrate that its compelling legitimate interest overrides the interests or the fundamental rights and freedoms of the data subject*". In other words, the exercise by an individual of its right to object to its personal data being processed by a controller essentially includes a balancing of rights and legitimate interests: on the one hand an individual is interested in having its data no longer processed and on the other hand a controller may have an interest in continuing to process such data despite the individuals' objections.

1.9. Security of personal data

Security of personal data is regulated under Section 2 of the GDPR and in particular under articles 32-34.

Article 32 deals with security of processing and sets the controller's and the processor's obligation to implement technical and organisational measures to ensure a level of security including among others:

a. the pseudonymisation and encryption of personal data;

b. the ability to ensure the ongoing confidentiality, integrity, availability and resilience of processing systems and services;

c. the ability to restore the availability and access to personal data in timely manner in the event of a physical or technical incident;

d. a process for regularly testing, assessing and evaluating the effectiveness of technical and organisational measures for ensuring the security of processing.

Another parameter of security of personal data is the notification of a personal data breach to the supervisory authority. Data Breach Notifications are regulated by article 33. A "personal data breach" is defined in the text of the GDPR, in Article 4(12), as "a breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorized disclosure of, or access to, personal data transmitted, stored or otherwise processed". When this happens, controllers shall, according to article 33, par. 1 "without undue delay and, where feasible, not later than 72 hours after having become aware of it, notify the personal data breach to the supervisory authority competent in accordance with Article 55, unless the personal data breach is unlikely to result in a risk to the rights and freedoms of natural persons. Where the notification to the supervisory authority is not made within 72 hours, it shall be accompanied by reasons for the delay".

The obligation of notification burdens the processor as well, who, shall notify the controller without undue delay after becoming aware of a personal data breach (article 33 par.2).

Paragraph 3 lists the minimum information the notification must contain, such as the nature of the data breach, the name and contact details of the data protection officer, the likely consequences of the personal data breach and the measures taken or proposed to be taken by the controller to address the personal data breach.

Finally, the communication of a data breach to the data subject is regulated under article 34. This obligation burdens the controller in any case where the personal data breach is likely to result in a high risk to the rights and freedoms of natural persons. The communication to the data subject shall describe in clear and plain language the nature of the personal data breach and contain at least the information and measures referred to in article 33(2). Paragraph 3 of article 34 sets the conditions under which the communication to the data subject is not required. In particular par. 3 reads as follows "*The communication to the data subject referred to in paragraph 1 shall not be required if any of the following conditions are met: a) the controller has implemented appropriate technical and organisational protection measures, and those measures were applied to the personal data affected by the personal data breach, in particular those that render the personal data unintelligible to any person who is not authorised to access it, such as encryption b) the controller has taken subsequent measures which ensure that the high risk to the rights and freedoms of data subjects referred to in paragraph 1 is no longer likely to materialise; c) it would involve disproportionate effort. In such a case, there shall instead be a public communication or similar measure whereby the data subjects are informed in an equally effective manner*".

2. Application of the GDPR in the NUTRISHIELD ETL process

2.1. Project's description and specifications

NUTRISHIELD aims at creating a personalised platform for the young. The platform will consist of novel methods & techniques, which analyse a wide range of biomarkers related to nutrition and health disorders. Based on findings, the platform then uses ICT, by expanding existing nutrition assistive mobile apps, in order to provide feedback and steering people towards a better nutrition.

Among the project's objectives are:

a) to deliver an innovative technology platform, the NUTRISHIELD Platform for personalised nutrition that people want to use and

b) to develop a set of new tools for assessing biomarkers relevant to personalised nutrition.

In this context NUTRISHIELD will develop a urine analyser, a human milk analyser, and a breath analyser, aiming at bringing analytical capabilities to a large number of practicing doctors, thus making personalised nutrition more widely adopted.

In more detail the NUTRISHIELD system consists of four major components:

a. the Measuring Devices

b. the Dashboard (a web application)

c. the Mobile Application and

d. the Backend System.

a. **The Measuring Devices** consist of three prototypes for urine, human milk and breath analysis. The operator, measuring with the prototypes, will measure as instructed and will add the information from the prototypes to the NUTRISHIELD platform. Beside the concentration values for each analyser, several additional info can be stored on the device, including measurement ID, Date & Time, Operator, Patient-ID/Sample ID, Absorbance and Status (Normal, Error, etc.).

b. **The Dashboard** is a web application used primarily by medical personnel to monitor patients, upload measurements and prepare the dietary and activity plans. Dashboard will act as the main web application which will be accessible by Administrators and Medical professionals. Dashboard will be the gateway to NUTRISHIELD Platform for Administrators (Data Managers and NUTRISHIELD Admins) and Medical professionals (Doctors & Lab Technicians). The Dashboard will be available to registered users only as a Web Portal/App

c. **The** NUTRISHIELD **Mobile Application** is used by patients/users for receiving notifications regarding their dietary and activity plan and logging the food consumed towards keeping a food journal. Running on a smartphone which must be connected to the internet, it interacts directly with the back-end, having a camera and enough storage to store meal consumption images and other information.

d. **The Back-end System** is a multi-functional component including many other sub-components such as databases, web services, the nutrition algorithm services and machine learning modules. It interacts with both the dashboard and the mobile application to provide the necessary support to their functionality.

2.2. The ETL processes in the NUTRISHIELD project

2.2.1. General

This report aims mainly to examine the process according to which any patient/research subjects'data collected by the NUTRISHIELD measuring devices, mentioned above, are transferred to the NUTRISHIELD dashboard. In other words, it will be elaborated how the communication between such sensors and the dashboard is conducted and what specific measures have already been undertaken or need to be implemented by the project's partners in order for such process to be compliant with the applicable legislation. Given in particular the sensitive nature of the data collected, health and/or genetic data in particular, it is essential that the NUTRISHIELD project complies with the GDPR and that the ETL processes are designed with the protection of personal data in mind.

It is also pointed out that the present report is the first version on the subject matter in question. A final version on data governance is scheduled for M24. This is of the essence since the ETL process needs to be further refined among the NUTRISHIELD partners. As a result, this analysis will focus on the process implemented at the time of drafting the present report and will provide the partners with useful guidance on compliance matters. At the same time, an effort will be made to present some basic guidelines in the event of alternative ETL processes is chosen for future use, in particular one where the sensors connect automatically to the NUTRISHIELD dashboard. Any final conclusions will be included in the final version D6.6.

2.2.2. Describing the ETL process: the role of the Healthcare Organisations

a. The ETL process applicable to NUTRISHIELD

Under the previous section a general description of the project's Measuring Devices was given. In more detail, the Measuring Devices of the NUTRISHIELD system are the Human Milk, Urine and Breath analyzers. Other than the measurements of the patient analytes (concentration values), and the analysis data, additional information could optionally be stored on each device, such as patient-ID (provided by medical personnel), date and time of measurement, measurement ID etc. (all these referred to as sensor data).

The Measuring Devices provide raw measurements, which can be either inputted directly in NUTRISHIELD by lab technicians or be analysed and the analysis results be inputted instead, if they are more meaningful. In order to better understand the process, it is important to note that the devices are used

in clinical or lab settings, they do not yet have IP connectivity, which is an issue to be finalised among project partners, and therefore for the moment are not directly connected to the NUTRISHIELD platform. In this context it has been decided among the partners that initially the input of measurement data that originate from the analyzers should occur by using dedicated dashboard input forms made available to the laboratory staff. Therefore, the operator measuring with the prototypes, will measure as instructed and will proceed to enter the results of the measurement to the platform using the dashboard.

This does not exclude the possibility of the prototype devices having network connectivity that allows the measurement data to be sent in an automatic manner. However, this is a possibility that lies outside of the main goals of the NUTRISHIELD project and will be evaluated in brief under this report for consistency reasons, as already mentioned above under 2.1.

b. How are Healthcare Organisations involved in the ETL processes?

Before addressing this question, it is noted that for the purposes of this report the term "Healthcare Organizations" will be used to denote the project partners that will participate as pilots in the NUTRISHIELD project. (the "**Healthcare Organisations"**)

To describe the ETL process in a few words, it is anticipated that any patient personal data collected by the sensors during the project's duration will be processed solely by the Healthcare Organizations' medical professionals and will be further inserted in the NUTRISHIELD dashboard manually by them (Doctors & Lab Technicians). Before the stage of collection, it is assumed that the Healthcare Organisations, and the medical professionals in particular, will inform the patients/participants that they participate in the NUTRISHIELD research, providing at the same time to them all the necessary information. At the stage of collection and processing of these personal data, it is the Healthcare Organisation's exclusive obligation, as data controller, to comply with the personal data processing principles provided for under the GDPR.

After collecting the personal data from the measuring devices, medical professionals should proceed with their anonymization. For this purpose, an abstract ID will be provided to each patient and this ID will replace any personally identifiable information, such as name, address, telephone number. This information will be kept separately at the Healthcare Organizations' care and responsibility. It is the Healthcare Organization's obligation to safeguard that all personal data that relate patients to this ID are kept in a separate patient management system already installed at the clinic. In any event it shall be made sure, at all times, during the project's execution, that all personal data will be anonymized before inserted in the NUTRISHIELD platform. The NUTRISHIELD platform, shall only store and process anonymous personal data.

What can be easily derived from the above analysis is that the role of the Healthcare Organizations and of the medical professionals in keeping the ETL process compliant with the GDPR is of crucial importance. For the present report, it is taken for granted that Healthcare Organisations participating in the NUTRISHIELD research have taken all the necessary measures for their internal GDPR compliance, as this is thoroughly described in the first part of this report. This however does not diminish the obligation of the NUTRISHIELD project to comply with the GDPR, as this obligation will be elaborated below.

2.3. The NUTRISHIELD platform and the processing of personal data

2.3.1. Are data introduced in the Nutrishield platform "personal data"?

The ETL process described above covers the uploading in the NUTRISHIELD Platform of the measurements derived from the Measurements devices. However, other data are expected to be introduced in the Platform. Some examples of these data are mentioned in the following paragraph. This section 2.3.1. examines the design and operation of the NUTRISHIELD platform as a medical tool for healthcare professionals and records the possible implications that may arise in connection to compliance with the GDPR.

It has already been made clear that medical professionals will anonymise any patients personal data before inserting them in the NUTRISHIELD platfrom. The process of anonymisation implemented by the Healthcare Organisations should be perceived as a measure adopted by NUTRISHIELD in order to comply with the principles of the GDPR. In more detail, the data that will be inserted in the NUTRISHIELD platform by the medical professionals, include among others, socio-demographic information, lifestyle information, dietary information and medical history of the patients/participants. It is noted that the list of categories of data may be enriched as the project progresses. As already stressed out, personal details such as name, email address, home address, telephone number will not be uploaded in the platform in order to safeguard that any connection of the measurements and other health or other personal data inserted in NUTRISHIELD, to the natural persons/patients, will be difficult, if not impossible. Upon patient registration completion, by the doctor, a patient ID will be generated. This will be the identity of the patient in the NUTRISHIELD platform. Considering that no patient personal data will be kept (name, surname, etc), the patient ID is the only "key" for the doctor to identify/search for a patient within NUTRISHIELD Database.

However, anonymasation of data, as this is conducted for the purposes of NUTRISHIELD, as well as implementation of the process described above, should not lead to the wrong conclusion that the data inserted in the platform are not personal data.

As explained in the first chapter of this report, personal data are any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person 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. According to recital 26 of the GDPR "*To determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, 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. 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 the data subject is not or no longer identifiable*".

By the combination of the above provisions, it is evident that the data inserted, in anonymised form, in the NUTRISHIELD **platform are personal data and most importantly are health data, namely**

special categories of personal data. Their anonymisation before their insertion in the platform does not automatically renders these data not personal. Keeping the patients' personal details, such as name, address etc. in separate files, makes the process of identification difficult but not impossible. It rests with the Healthcare Organizations participating in the project's pilots to warrant that these data will remain safe and that the provisions of the GDPR will be respected. However, it should also be made sure that the NUTRISHIELD Project applies the GDPR principles, as this shall be further examined in the sections that follows.

2.3.2. Application of the personal data processing principles to NUTRISHIELD

a. Lawfulness of processing and informed consent in particular

In order to address this issue, it should first be noted that in the NUTRISHIELD project, the Healthcare Organizations participating in the research, act as "controllers", which means that, as far as their processing activities that take place in the context of NUTRISHIELD are concerned, it lies with them to comply with the GDPR principles. In this context, the Healthcare Organizations should make sure that the processing is lawful in the sense that they apply their internal GDPR- compliance policies, in particular they have acquired the necessary informed consent by their patients/research participants, as per the requirements imposed by their national legislation (see par 1.7).

Healthcare Organizations fall under the exceptions provided for under article 9.2 (h) of the GDPR, as regards the processing of special categories of data, according to which, first processing of special categories of personal data is permitted and, second, this can be done without the patient's consent.

In more detail article 9.2 (h) states that processing of special categories of data shall not be prohibited when "*processing is necessary for the purposes of preventive or occupational medicine, for the assessment of the working capacity of the employee, medical diagnosis, the provision of health or social care or treatment or the management of health or social care systems and services on the basis of Union or Member State law or pursuant to contract with a health professional and subject to the conditions and safeguards referred to in paragraph 3*". **As a result, lawful processing of patients' sensitive personal data by Healthcare Organisations is permitted upon this legal basis, subject also to any national law requirements.**

However, for their participation in the Nutrishield research, it is advised that Healthcare Organisations acquire informed consent for processing their patients' personal data for the specific research purpose. The reason for that is that processing for research purposes can be considered further processing (Art. 6.4.) and does not fall under the purpose of article 9.2 (h). In other words, Healthcare Organizations should apply both provisions when processing personal data in the context of NUTRISHIELD. **It is also noted, that, as indicated in the legal analysis of the first chapter, Healthcare Organizations need to acquire two sets of informed consent from their patients, namely one for their participation in research from an ethics perspective, and one for the processing of their personal data for this specific purpose. This, evidently, does not preclude both consents to be incorporated into a single form. Extra precautions should be undertaken as far as children participating in the clinical trials are concerned. In this case, depending on the child's exact age, both consent forms should be signed by the legally authorised representative of the child's.**

In the same context, it is useful to include in this report the recent opinion issued by the European Data Protection Board on the interplay between the Clinical Trials Regulation (CTR)[9] and the GDPR. The opinion focuses on the applicable legal basis for the processing of personal data in the course of a clinical trial protocol. Its main conclusion is that the informed consent foreseen under the CTR (article 29) must not be confused with the notion of consent as a legal ground for the processing of personal data under the GDPR. Therefore, the organisation conducting the clinical trial should always have in place both forms for informed consent signed and updated by the patients/participants in the clinical trials[10].

b. The principle of transparency

As far as the principle of transparency is concerned, article 12(1) of the GDPR states that it is the controller's obligation to warrant transparency of processing. According to recital 39 of the Regulation, the principle of transparency requires that any information and communication relating to the processing of those personal data be easily accessible and easy to understand, and that clear and plain language be used. That principle concerns, in particular, information to the data subjects on the identity of the controller and the purposes of the processing and further information to ensure fair and transparent processing in respect of the natural persons concerned and their right to obtain confirmation and communication of personal data concerning them which are being processed. The Healthcare Organisations participating in the NUTRISHIELD research should, for any data processing that will be conducted for the NUTRISHIELD purposes, apply the principle of transparency. They should in particular provide all information included in articles 13 and 14 of the GDPR, such as the identity and the contact details of the controller, the contact details of their DPO, the purposes of the processing, the existence of the rights provided to data subjects under the GDPR etc.

c. The principle of accountability

Furthermore, Healthcare Organisations and health care infrastructures participating in the NUTRISHIELD research should be able to demonstrate compliance with the GDPR. In this context, they should always apply the necessary measures within their organizations in order to be able to demonstrate both to the individuals concerned and to any future controls by the Data Protection Authorities that the data protection legislation has been observed. Indicatively, it is anticipated that they apply an updated internal data protection policy at all times and that they have an appointed DPO.

d. Purpose limitation and data minimisation

Anonymisation of patients' personal data, before their insertion in the NUTRISHIELD platform, fulfils, to the best extent possible, the principle of purpose limitation and data minimisation. By the process of

[9] See footnote 6

[10] See Opinion 3/2019 concerning the Questions and Answers on the interplay between the Clinical trials Regulation (CTR) and the General Data Protection regulation (GDPR), adopted by the EDPB on 23 January 2019

https://www.dataprotection.ro/servlet/ViewDocument?id=1629

anonymisation, it is made certain that technical and organisational measures have been implemented in order to safeguard these principles. It is mostly safeguarded that any personal data will be processed to the extent absolutely necessary for the purposes of the project and solely for such purpose.

2.3.3. Data subjects' rights

The rights the GDPR provides for the data subjects are elaborated in the relative section of the first part of this report. As regards the exercise of these in the context of the NUTRISHIELD project, it is expected that all such rights will be respected throughout the project's duration. In more detail, the right to information and access to personal data the right to erasure and right to object should be safeguarded. Again, given that the personal data of the patients participating in NUTRISHIELD will be collected, stored and processed by the Healthcare Organisations, it is them that need to enable the data subjects in exercising these rights.

2.3.4. Security of the processing in the context of the NUTRISHIELD project

Whenever and however personal data are being collected in the context of a project, it is the researchers ethical and legal obligation to ensure that participants' information is properly protected. Security of processing is regulated under section 2 of the GDPR and in particular under articles 32-34. Article 32 lists the measures the controller and the processor shall implement in order to ensure a level of security appropriate to the risk. These measures include among others:

a) the pseudonymisation and encryption of personal data;

b) the ability to ensure the ongoing confidentiality, integrity, availability and resilience of processing systems and services;

c) the ability to restore the availability and access to personal data in a timely manner in the event of a physical or technical incident;

d) a process for regularly testing, assessing, and evaluating the effectiveness of technical and organisational measures for ensuring the security of the processing.

Security of the processing is further strengthened with the process of notifying a personal data breach to the supervisory authority, as this is described in article 33 of the GDPR. Finally, security is completed with the process of article 34 of the Regulation, namely the communication of a personal data breach to the data subject.

The GDPR requires all data controllers and processors to implement appropriate technical and organisational measures to ensure a level of data security that is commensurate to the risks faced by the data subjects in the event of unauthorised access to, or disclosure, accidental deletion or destruction of, their data (art.32 GDPR).

In NUTRISHIELD, the protection of personal data that are processed during its execution and the obligation to ensure security of the processing is further strengthened by the fact that the personal data in question are

health and potentially genetic data. It has already been established in the previous sections that Healthcare Organisations, participating in NUTRISHIELD, are, as controllers, burdened with the respective obligations. However, the NUTRISHIELD project, as data processor, must comply with the obligations provided by the GDPR, including of course the security of processing. The technical and organisational measures implemented by the Healthcare Organisations is beyond the scope of this report, nevertheless the compliance of the project with the GDPR relies to a great extent to these Healthcare Organisations' internal GDPR-compliance exercise.

NUTRISHIELD on its turn, should provide details of the technical and organisational measures that will be implemented to protect the personal data processed in the course of the NUTRISHIELD research. Such measures may include, the anonymisation process (as this was elaborated above), the pseudonymisation and encryption of personal data, and policies and procedures to ensure the confidentiality, integrity, availability and resilience of the NUTRISHIELD processing systems.

The Commission has published a list with 10 do's and don'ts that research participants should have in mind when it comes to data security[11].

	DO's	DON'Ts
1	use GDPR-compliant tools to collect, process and store research subjects' personal data;	collect data on a personal device such as a smartphone without ensuring that they are properly protected (e.g. consider the implications of automatic back-ups to the cloud, and the device's security features);
2	take communications security seriously, and devise and implement dedicated protocols for your project as necessary;	use free services that may use your participants' data for their own purposes in lieu of payment, or collect data or communicate with research participants via social media platforms without first assessing the data protection implications;
3	check the terms and conditions of all of the service providers you use (software, applications, storage, etc.) to process personal data within your project, in order to identify and mitigate risks to the data subjects	use unencrypted email, SMS or insecure 'voice over IP' platforms to communicate with vulnerable participants or those who may be subject to state surveillance;

[11] See Ethics and Data Protection, European Commission, November 2018, https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/ethics/h2020_hi_ethics-data-protection_en.pdf

4	encrypt your research data and/or the devices on which they are stored, and ensure that keys/passwords are appropriately protected; and	expose personal data to unauthorised access or use when accessing them remotely (e.g. by using insecure wifi connections) or travelling to countries where your devices may be inspected or seized; and
5	consult your DPO or a suitably qualified expert for advice on how to achieve a level of data security that is commensurate to the risks to your data subjects.	assume that your research partners, collaborators or service providers have appropriate information security and data protection policies without checking that this is the case.

2.3.5. The NUTRISHIELD platform and the principles of data protection by design and by default

a. What is data protection by design and by default?

The GDPR, for the first time, addresses data protection by design as a legal obligation for data controllers and processors. Furthermore, it introduces the obligation of data protection by default and establishes the protection of personal data as a default property of systems and services.

Data protection by design aims to build data protection and privacy into the design of processing operations and information systems, in order to comply with the data protection principles. In other words, it aims to prevent privacy invasive events before they happen.

Recital 78 of the GDPR states that the controller should adopt internal policies and implement measures which meet in particular the principles of data protection by design and data protection by default. These principles are regulated under article 25 of the Regulation which reads as follows: "*Taking into account the state of the art, the cost of implementation and the nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for rights and freedoms of natural persons posed by the processing, the controller shall, both at the time of the determination of the means for processing and at the time of the processing itself, 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 necessary safeguards into the processing in order to meet the requirements of this Regulation and protect the rights of data subjects*".

As referred to in the Guidelines of the European Data Protection Board on data protection by design and by default [12], a technical or organisational measure can be anything form the use of advanced technical solutions to the basic training of the personnel. There is no requirement as to the sophistication of a measure, as long as it is appropriate for implementing the data protection principles effectively and proportionately to the risk.

Data protection by default, in a nutshell, means that the user profile settings must be automatically data protection friendly. Thus, the default settings must be designed with data protection in mind. Data protection by default is regulated under par. 2 of article 25 which reads as follows:

"*The controller shall 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. That obligation applies to the amount of personal data collected, the extent of their processing, the period of their storage and their accessibility. In particular, such measures shall ensure that by default personal data are not made accessible without the individual's intervention to an indefinite number of natural persons".*

As regards the obligation of data protection by default, the EDPS in its preliminary opinion on privacy by design[13], mentions, among others that: "*Following the application of the principle of data protection by design, organisations must, by default, only process personal data necessary for each specific purpose defined in compliance with the law and transparently notified to the individuals concerned. While it can be argued that this obligation is already implicit in the "purpose limitation" and "data minimisation" principles in both the design and operation phases, the explicit rule stresses the importance of taking technical measures to meet the expectations of the individuals whose data are processed, not to have their data processed for other purposes than what the product and service is basically and strictly meant to do, leaving by default any further use turned off, for instance through configuration settings*".

b. How to apply the principles of data protection by design and by default in the NUTRISHIELD platform.

As already demonstrated above, it is anticipated that the NUTRISHIELD platform will collect and process personal data (in anonymised form, but still personal data for GDPR purposes). At this stage of the NUTRISHIELD research, the platform is designed as a tool that is addressed exclusively to medical professionals and not to users/patients. This however does not mean that this is not an option that needs to be taken into consideration when designing the platform. At the same time, it was clarified in section 2.2.2 (a), where the ETL process is described, that, at the time of writing this report, the partners have not yet finalised how the NUTRISHIELD sensors will communicate with the NUTRISHIELD dashboard. At the time being this involves a manual procedure, without however the possibility of the prototype devices

[12] See EDPB Guidelines 4/2019 on Article 25 Data Protection by Design and by Default, November 2019.
https://edpb.europa.eu/sites/edpb/files/consultation/edpb_guidelines_201904_dataprotection_by_design_and_by_default.pdf

[13] Preliminary Opinion 5/2018 on Privacy by design by the EDPS, published on 31 May 2018

having network connectivity that allows the measurement data to be sent in an automatic manner to the dashboard, being excluded.

Regardless of the use of the platform as a professional tool or as a user-friendly tool, and regardless of the method used for the collection of data- automatic process or not- it is undisputable that the principle of data protection by design and by default must apply. The NUTRISHIELD platform should, at the early stage of its "architectural building", be designed in such a way that any personal data processing will be conducted with respect to data subjects' rights and freedoms. In other words, privacy measures and privacy enhancing technologies (PETs)[14] should be directly embedded into the design of the NUTRISHIELD application.

Measures to achieve data protection by design could include:

• the pseudonymisation or anonymisation of personal data;

• data minimisation

• applied cryptography (e.g. encryption and hashing)

• using data-protection focused service providers and storage platforms; and

• arrangements that enable data subjects to exercise their fundamental rights (e.g. as regards direct access to their personal data and consent to its use or transfer).

According to the European Data Protection Board's guidelines on article 25[15], key design and default elements may include:

In terms of transparency

Clarity – Information shall be in clear and plain language, concise and intelligible.
Semantics – Communication shall have a clear meaning to the audience in question
Accessibility - Information shall be easily accessible for the data subject.
Contextual – Information shall be provided at the relevant time and in the appropriate form
Relevance – Information shall be relevant and applicable to the specific data subject
Universal design – Information shall be accessible to all, include use of machine-readable languages to facilitate and automate readability and clarity
Comprehensible – Data subjects shall have a fair understanding of what they can expect with regards to the processing of their personal data, particularly when the data subjects are children or other vulnerable groups
Multi-channel – Information should be provided in different channels and media, beyond the textual, to increase the probability for the information to effectively reach the data subject

In terms of lawfulness of processing

Relevance – The correct legal basis shall be applied to the processing
Differentiation – The controller shall differentiate between the legal basis used for each processing activity

[14] Privacy Enhancing Technologies, i.e. specific technological solutions to certain privacy related issues in systems design, have preceded the idea of a comprehensive privacy engineering approach and today they can be considered as quality basic building blocks for engineering privacy

[15] See Footnote 10

Specified purpose - The appropriate legal basis must be clearly connected to the specific purpose of processing.
Necessary – Processing must be necessary for the purpose to be lawful. It is an objective test which involves an objective assessment of realistic alternatives of achieving the purpose.
Autonomy – The data subject should be granted the highest degree of autonomy as possible with respect to control over personal data.
Consent withdrawal – The processing shall facilitate withdrawal of consent. Withdrawal shall be as easy as giving consent. If not, any given consent is not valid.
Balancing of interests – Where legitimate interests is the legal basis, the controller must carry out an objectively weighted balancing of interests. There shall be measures and safeguards to mitigate the negative impact on the data subjects, and the controller should disclose their assessment of the balancing of interests
Predetermination – The legal basis shall be established before the processing takes place.
Cessation – If the legal basis ceases to apply, the processing shall cease accordingly
Adjust – If there is a valid change of legal basis for the processing, the actual processing must be adjusted in accordance with the new legal basis
Default configurations – Processing must be limited to what the legal basis strictly gives grounds for.
Allocation of responsibility – Whenever joint controllership is envisaged, the parties must apportion in a clear and transparent way their respective responsibilities vis-à-vis the data subject

In terms of purpose limitation

Predetermination – The legitimate purposes must be determined before the design of the processing.
Specificity – The purposes must be specific to the processing and make it explicitly clear why personal data is being processed.
Purpose orientation – The purpose of processing should guide the design of the processing and set processing boundaries.
Necessity – The purpose determines what personal data is necessary for the processing
Compatibility – Any new purpose must be compatible with the original purpose for which the data was collected and guide relevant changes in design.
Limit further processing – The controller should not connect datasets or perform any further processing for new incompatible purposes.
Review – The controller must regularly review whether the processing is necessary for the purposes for which the data was collected and test the design against purpose limitation.
Technical limitations of reuse – The controller should use technical measures, including hashing and cryptography, to limit the possibility of repurposing personal data.

In terms of data minimisation

Data avoidance - Avoid processing personal data altogether when this is possible for the relevant purpose
Relevance – Personal data shall be relevant to the processing in question, and the controller shall be able to demonstrate this relevance.

Necessity – Each personal data element shall be necessary for the specified purposes and should only be processed if it is not possible to fulfil the purpose by other means.
Limitation – Limit the amount of personal data collected to what is necessary for the purpose
Aggregation – Use aggregated data when possible.
Pseudonymization – Pseudonymize personal data as soon as it is no longer necessary to have directly identifiable personal data, and store identification keys separately.
Anonymization and deletion – Where personal data is not, or no longer necessary for the purpose, personal data shall be anonymized or deleted.
Data flow – The data flow shall be made efficient enough to not create more copies, or entry points for data collection than necessary.
"State of the art" – The controller should apply available and suitable technologies for data avoidance and minimisation.

The principles of data protection by design and by default should be the main concern of the NUTRISHIELD partners. It was clarified in section 2.1., where the ETL process is described, that, at the time of writing this report, the partners have not yet finalised how the NUTRISHIELD sensors will communicate with the NUTRISHIELD dashboard. Regardless however of the method that will be used for the uploading of the data in the NUTRISHIELD Platform (automatically or not), the technical and organisational measures suggested herein should be taken into consideration.

3. Conclusion

The present report tries to bring to the surface the main issues that relate to the personal data processing activities that are anticipated to take place in the context of the NUTRISHIELD project. In this context, the ETL process is closely examined and specific guidance is provided to the partners in order to implement this procedure according to the main GDPR principles. At the same time great attention is given to the collection and processing of personal data by the NUTRISHIELD Platform. The nature of these data as health and genetic data makes, compliance of the Platform and the project in its total with the GDPR, even more critical. To this end some suggestions and advice are provided to the project partners in an effort to facilitate them during the execution of the project in order to be always in line with applicable personal data protection legislation. Finally, it is pointed out that this report is a first version on the subject matter in question and should be viewed as a preliminary approach regarding the project's conformity with the GDPR. The final remarks and suggestions will be included in the final version, that is due on month 24, following the project's progress.

4. References

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 personal data and on the free movement of such data, and repealing Directive 95/46/EC.

Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data (replaced by the GDPR).

A Preliminary Opinion on data protection and scientific research, by the EDPO, 6.1.2020 https://edps.europa.eu/sites/edp/files/publication/20-01-06_opinion_research_en.pdf

Regulation (EU) No 1291/2013 of the European Parliament and of the Council of 11 December 2013 establishing Horizon 2020 - the Framework Programme for Research and Innovation (2014-2020) and repealing Decision No 1982/2006/EC.

EDPB, Guidelines 4/2019 on Article 25 Data Protection by Design and by Default, November 2019 https://edpb.europa.eu/sites/edpb/files/consultation/edpb_guidelines_201904_dataprotection_by_design_and_by_default.pdf

Article 29 Working Party Guidelines on consent under Regulation 2016/679 adopted on 28 November 2017 https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=623051

See Opinion 3/2019 concerning the Questions and Answers on the interplay between the Clinical trials Regulation (CTR) and the General Data Protection regulation (GDPR), adopted by the EDPB on 23 January 2019 https://www.dataprotection.ro/servlet/ViewDocument?id=1629