D8.1 POPD – Requirement

Flying Forward 2020 is funded by the European Union H2020 Research and Innovation Programme under Grant Agreement No. 101006828

Contents

Document Identification

Authors	Anneke Veldstra (BRAIN), Ruud van Iwaarden (BRAIN), Ingrid van Vught
Internal reviewers	n.a.
Work package	WP8 “Ethics”
Task(s) involved	n.a.
Nature	R : Document, report
Dissemination level	PU : Public

Version	Date	Contributor	Description
V0.01	03/06/2021	Anneke Veldstra (BRAIN)	Initial draft
V0.02	15/11/2021	Ruud van Iwaarden (BRAIN), Anneke Veldstra (BRAIN)	Revision
V1	14/05/2022	Anneke Veldstra (BRAIN)	Final version V1
V2	27/09/2022	Anneke Veldstra (BRAIN), Ruud van Iwaarden (BRAIN)	Revised and adapted to V2

Acronyms and Abbreviations

Acronym	Title
CA	Consortium Agreement
DMP	Data Management Plan
GDPR	General Data Protection Regulation
FAIR	Findable, Accessible, Interoperable, Reusable
POPD	Protection of Personal Data
FF2020	Flying Forward 2020

References

Reference	Name
[RD1]	FF2020 project evaluation – Ethics Summary Report
[RD2]	FF2020 project consortium agreement

1. Introduction

Since the FF2020 research and innovation action envisages engagement of various stakeholders and citizen through solution's architectures design by co-creation, operational concept definition, and end-users' requirements, it qualifies as a research that involves collection and processing of personal data.

At another level, by conducting close range electro/optical and radar imaging of the urban surroundings in which human inhabitants reside, possibility of identifying individuals exists and as such it qualifies as collecting of personal data with potential of being used for person's identification.

Thus, the implementation of the General Data Protection Regulation 2016/679 is required. According to this Regulation, consortium shall have established data controllers and processors which are fully accountable for the data processing operations.

The outcome of the Ethics evaluation of the FF2020 project proposal obliged the consortium to define aforementioned procedures and forms that will be used throughout the project lifetime. This post-grant obligation in Ethics Summary Report is titled as Protection of Personal Data – Requirement.

With respect to the ethics issues concerning Protection of personal data as specified in the [RD1] Ethics Summary Report, bellow response by FF2020 project consortium applies:

[…] *Does this research involve personal data collection and/or processing? Yes* […]

Response by FF2020

Personal data will be collected and processed during the project activities, i.e. investigation of citizens' engagement and perspectives via interviews and workshops.

[…] Ethics Issues

Ethics recommendations

Although no ethics issues have been recognised in Part A or Part B

The beneficiary is reminded that under the General Data Protection Regulation 2016/679, the data controllers and processors are fully accountable for the data processing operations. Any violation of the data subject rights may lead to sanctions as described in Chapter VIII, art.77-84.

Ethics Opinion

Conditional ethics clearance (i.e. clearance is subject to conditions, i.e. ethics requirements. The requirements must either be fulfilled before grant signature or become part of the grant agreement). […]

2. Personal data collection and/or processing

2.1 Relevance of the data collection and/or processing for the project

In this document project consortium explains how all of the data it intends to process is relevant and limited to the purposes of the research project (in accordance with the 'data minimisation 'principle).

Two levels of personal data collection and/or processing exists within FF2020 project. Albeit not serving for identifying or profiling an individual, it is rather considered as having a potential for being used without authorisation or against the General Data Protection Regulation.

At one level, by collecting the information from citizens about their preferences in relation to questionnaires used in FF2020 have a potential for unauthorized use if metadata inappropriately safeguarded.

At another level, by visual or radar imaging of urban surroundings in which humans reside, potential for identification of individual exists if imaging data are not managed in a secure way (including storage and access during and after the project lifetime)..

2.2 Data minimization principle

The data minimisation principle is expressed in Article 5(1)(c) of the GDPR and Article 4(1)(c) of Regulation (EU) 2018/1725, which provide that personal data must be "adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed".

The principle of "data minimisation" within FF2020 is fully aligned with the above-mentioned regulations and comprehends that a data controller should limit the collection of personal information to what is directly relevant and necessary to accomplish a specified purpose. As FF2020 considers two levels of personal data collection and/or processing, a data controller is foreseen for each type of considered data:

- For the information collected from citizens about their preferences in relation to questionnaires used in FF2020, the responsible data controller is Full name and email address.
- For visual imaging of urban surroundings in which humans reside, the responsible data controller is Full name and email address.
- For radar imaging of urban surroundings in which humans reside, data controller is Full name and email address.

The data controllers make sure that data are retain for as long as is necessary to fulfil their purpose within FF2020 activities (as described in the GA). In other words, data controllers take care to collect only the personal data they really need, and should keep it only for as long as they need it.

Personal data will be fully anonymized or pseudonymised wherever possible and securely stored with access possible only to the strictly necessary authorised persons. When possible, FF2020 experiments will be populated with fictional data, as well as data on user experiences collected in the experimentations will be anonymised at source. The confidentiality and

transparency of the process will be guaranteed, as well as the treatment of all research participants with respect at all times, with proper protection of their anonymity. All personal data will be used solely by the project and only for testing and validation purposes, without providing them to subjects not directly involved in the experiment concerned.

3. Technical and organizational measures for safeguard the rights and freedoms of the data subjects/research partners

Technical and organisational measures for safeguard the rights and freedoms of the data subjects/research participants include defined procedures for citizens data collection (defined in more details in FF2020 D8.1) and technical measures described in FF2020-DMP.

These, among others involve, the procedures and criteria that will be used to identify and or recruit participants. In FF2020 the selection of the participants is only foreseen within the scope of the citizens' workshops, and where practicably reasonable, selection of the groups of citizens aimed at participating in aforementioned workshops is made with regard to balancing among the gender, geographical, age and educational representation.

Furthermore, in order to comply with ethics principles and data privacy regulations, put in place by the FF2020 consortium and the European Commission, we require each participant to visit the page FF2020 website. There, participant can read the Information sheet explaining the terms of their engagement in consultation, and fill in the forms at the above link. Also, participant will find a quick and guided procedure needed to comply for GDPR and ethics duties, allowing FF2020 consortium to start participant's involvement in project activities.

Workflow for stakeholder compliance with ethics principles and privacy regulations is defined as following:

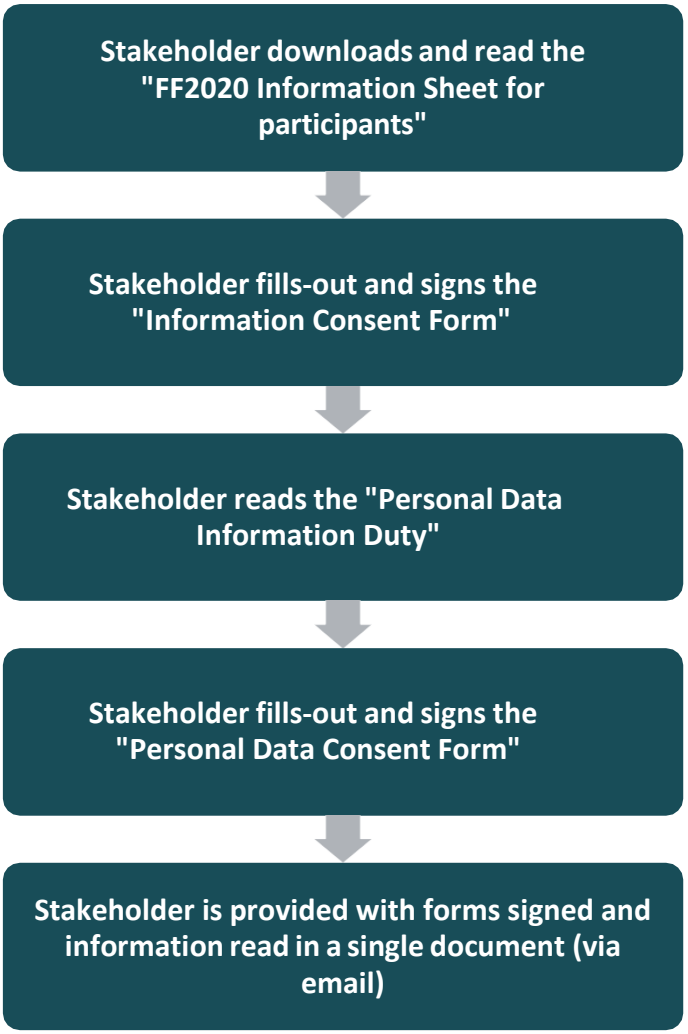

Figure 1 Workflow 1 - Compliance with ethics principles and privacy regulations

The personal data management in FF2020 project activities are delegated to the following roles and persons:

a) Task Coordinator:

In the FF2020 project, the parties have a common goal that they have defined and will pursue. The Task Coordinator is obliged to address the topic concerning privacy issues to the parties responsible for executing the Task beforehand. Each party, however, is responsible to comply with the GDPR following its own privacy policy and protocols. For work package D8.1 Ruud van Iwaarden is the Task Coordinator, his email address is r.vaniwaarden@brainportdevelopment.nl

b) Consortium Coordinator's Data Protection Officer:

Brainport Development as Consortium Coordinator is not obliged to appoint a Data Protection Officer according to art. 37 GDPR. Brainport Development, however, has appointed a Privacy Office and has a privacy policy including a data breach protocol. This privacy policy including its additional protocols is available at the Privacy Office of the Consortium Coordinator and will be followed by the Consortium Coordinator for the execution of its own tasks as set out in the project plan as well as for privacy issues on a consortium level.

The procedures concerning privacy issues of the Consortium Coordinator are to be found in:

a. Privacy policy
b. Privacy statement
c. Data breach procedure
d. Procedure Rights of data subjects
e. General Processing Register
f. Retention policy (on-going)

The Privacy Officer can be reached through email: privacy@brainportdevelopment.nl. The full form is also included as an Annex contained in a dedicated deliverable (FF2020-DMP). The collected data are stored in a secured data environment, in line with the FF2020 and EU Horizon 2020 safety requirements and as described in the FF2020's DMP.

4. Security measures

Security matters that will be implemented to prevent unauthorized access to personal data or the equipment used for processing

FF2020 security measures include mandatory authorisation to access data. Only project partners are authorised to access the data shared space. Next to this, data owner can additionally limit access to data by sharing it only with a limited group of project participants when this is justified and in line with data minimisation principle.

Furthermore, data are always stored on EU server, while the data recovery and backup are in charge of the shared space provider, where backups are performed in two parts (files and database of content).

5. Anonymisation/pseudo-anonymisation techniques

Within FF2020, data anonymisation is comprehended as a group of techniques used to anonymise the data. For data to be truly anonymised we consider that they must be processed irreversibly in such a manner that it can no longer be used to identify an individual by using “all the means likely reasonably to be used” as defined by European commission. Furthermore, within FF2020, we consider two main approaches to anonymisation, randomisation and generalisation, where the most suitable on for the data set at hand will be selected by data controller and implemented by the dedicated team member within the data owners’ institution.

In this context, randomisation refers to a group of techniques that alter the veracity of the data so that the data can no longer be referred to a specific individual. Randomisation techniques considered within FF2020, among others, include noise addition and permutation.

The noise addition technique consists of intentionally introducing the noise in the dataset prior to publication by modifying features, such that they become less accurate whilst retaining the overall distribution. When processing such anonymised dataset, an analyst assumes that values are accurate, however, this will only be correct to a certain degree. For instance, if workshop participant provided details about his home town, the anonymised dataset may contain modified location precision that is accurate to only several kilometres in latitude and longitude. This process should make a home location of an individual, and thus the individual himself, not identifiable from the dataset. Hence, if this randomisation strategy is applied effectively, an independent observer should not be able to identify individual nor detect the noise addition principles and, respectively, should not be able to repair the data, to their original values.

Permutation technique consists of disarranging the values of features by artificially associating part of them to different individual in the dataset. This technique is particularly convenient when it is relevant to retain the exact distribution of each feature within the dataset. Alike noise addition technique, permutation may not provide anonymisation by itself and should always be combined with the removal of obvious features and quasi-identifiers.

The differential privacy is not utilised within FF2020 as it is used when anonymised views of a dataset need to be created, typically upon the request of an authorised third party and through a subset of queries. However, this will not be the case within FF2020 as data will not be shared with third parties, as described in the FF2020's DMP.

Generalisation is the second main strategy among anonymisation techniques that is considered within FF2020, and it consists of generalising the features of individuals by modifying the respective scale or creating a broader categorisation. For instance, generalising home location from town level precision to a country level or using a week as time reference rather than a day. Albeit generalisation can be effective to prevent a possibility to isolate some or all records which identify an individual in the dataset (singling out), it necessitates dedicated and sophisticated quantitative approaches to avert likability (ability to link records concerning the same individual, or a group of individuals, either in the same database or across different databases) and inference (possibility to deduce, with significant probability, the value of a feature from other features). Some of the generalisation techniques considered within FF2020 include aggregation and k-anonymity with their extensions as L-diversity and T-closeness.

Aggregation and k-anonymity techniques assures that any individual in the dataset is indistinguishable from at least other k−1 individuals in terms of quasi-identifier attributes values. To achieve this, the feature values are aggregated to an extent such that everyone in that group shares the same value.

L-diversity is a refinement to the k-anonymity that aims to ensure that in each equivalence class every attribute has at least l different values well represented. In such cases, a potential attacker with existing background familiarity on a specific individual would remain with significant uncertainty due to the confinement of the occurrence of equivalence classes with poor feature variability.

T-closeness is a further extension of l-diversity that ensures that the distance between the distribution of a sensitive feature in equivalent class and the initial distribution of the feature is no more than a threshold t. This technique is convenient when it is relevant to keep the data as close as possible to the original one.

Pseudonymisation is a different type of technique, closely related to the anonymisation that focuses on reducing linkability of a dataset with the original identity of an individual. This is achieved by substituting a recognisable, or unique, feature in a dataset by another (pseudo) feature. For instance, replacing personal names in the dataset by unique numerical identifiers as a pseudo feature. The result of this process is not an anonymised dataset as it is likely to still allow indirect identification of an individual. Hence, it is used within FF2020 only as a useful security measure in combination with the above described anonymisation techniques.

6. GDPR implementation in the project

The GDPR are implemented as a set of procedures contained in a dedicated deliverable.

Flying Forward 2020 is funded by the European Union H2020 Research and Innovation Programme under Grant Agreement No. 101006828