Plan Overview

A Data Management Plan created using DMPonline

Title: Flexigrobots

Creator: Fernando Aguilar

Data Manager: Fernando Aguilar

Contributor: Anil Turkmayali

Affiliation: Other

Funder: European Commission

Template: Horizon 2020 DMP

Project abstract:

In an effort to enhance agricultural production, agricultural robots are being increasingly adopted. The number of discreet tasks to be automated, however, significantly reduces the flexibility of current solutions, thus impacting efficiency and limiting their take-up. The EU-funded FLEXIGROBOTS project aims to develop a platform that will help build heterogeneous multi-robot systems, allowing for vastly improved flexibility by using existing robots for multiple tasks. This could help produce higher-value data from a variety of sources and sensors, increasing operational autonomy and precision while greatly reducing costs and encouraging investment in robotics.

ID: 136468

Start date: 01-01-2021

End date: 31-12-2023

Last modified: 04-12-2023

Grant number / URL: 101017111

Flexigrobots - Initial DMP

1. Data summary

Provide a summary of the data addressing the following issues:

- **State the purpose of the data collection/generation**
- **Explain the relation to the objectives of the project**
- **Specify the types and formats of data generated/collected**
- **Specify if existing data is being re-used (if any)**
- **Specify the origin of the data**
- **State the expected size of the data (if known)**
- **Outline the data utility: to whom will it be useful**

More and more data is becoming publicly available through government open data policies and the open publishing of research data, while simultaneously satellite programmes like Copernicus are producing a mass of data. In the course of the project, a set of common data services to enable data collection and integration, data publication and sharing, and data security and trust, following the International Data Space Association (IDSA) reference architecture, enabling pilots to share data but also with/from third parties, will be integrated into the FlexiGroBots platform. During the lifecycle of the FlexiGroBots project, various data sets will be collected. This includes large data sets consisting of a wide range of data types (relational, text, multi-structured data, images, etc.) from numerous sources.

FlexiGroBots aims at enabling an agricultural data economy for farmers to profit directly, by selling operational data gathered in their fields, and indirectly, by acquiring third-party datasets and/or machine learning (ML) models for infusing customised artificial intelligence (AI) into their systems from day one. An important capability of the FlexiGroBots platform would an Agricultural Data Space which will allow farmers to share data of their robotics operations in their fields, with their crops, with stakeholders downstream the Agri-Food value chain. This data would be especially beneficial for the processing and distribution phases, including retailers, for quality management, forecasting, and recommendations of their respective operations. Furthermore, the open nature of the FlexiGroBots platform will enable other agricultural sectors (e.g. crops) to use it and benefit from the capabilities that were initially validated for the three FlexiGroBots pilots, for the grapevines, rapeseed and blueberry.

2. FAIR data

2.1 Making data findable, including provisions for metadata:

- **Outline the discoverability of data (metadata provision)**
- **Outline the identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers?**
- **Outline naming conventions used**
- **Outline the approach towards search keyword**
- **Outline the approach for clear versioning**
- **Specify standards for metadata creation (if any). If there are no standards in your discipline describe what metadata will be created and how**

During the implementation of the project, different data will be obtained and generated in relation to each work package, task and pilot and presented in the corresponding deliverables. All the partners have access to documents and deliverables in OwnCloud, where different folders have been created to make access to information easier. One of the folders is a repository for the deliverables, both in .docx and .pdf formats.

For datasets, we will have a designated folder where all the questionnaires will be collected, according to the iterations (initial and M06, periodic at M17 and final at M35). In order to be able to distinguish and easily identify datasets, each data set will be assigned a unique name (Name of dataset). This name can also be used as the identifier of the data sets. All data files produced, including emails, include the term “FlexiGroBots”, followed by file name which briefly describes its content and task number, e.g. “FlexiGroBots_PersonalOpinionData_Task6.1”.

In addition to this, a summary document in form of a datasets catalogue will be prepared with the main purpose to ease the identification of datasets generated in the course of the project. This catalogue will be a “living” document and constantly updated during the lifetime of the project.

Metadata is data on the research data itself, which enables researchers to find suitable data in an online repository. FlexiGroBots will provide metadata in the suitable standardized formats requested by the repositories used.

As required by Article 29.2 of the Grant Agreement, the bibliographic metadata must be in a standard format and must include all of the following:

- the terms “European Union (EU)” and “Horizon 2020”;
- the name of the action, acronym and grant number;
- the publication date, and length of embargo period if applicable, and
- a persistent identifier.

For research data to be found and subsequently reused, it is essential to provide a detailed and meaningful description in the metadata.

Datasets that will be made publicly available might be uploaded to open repositories like Zenodo. Zenodo is a general-purpose open-access repository developed under the European OpenAIRE program and operated by CERN. Zenodo (https://zenodo.org) is an open repository for all scholarships, enabling researchers from all disciplines to share and preserve their research outputs, regardless of size or format. Free to upload and free to access, Zenodo makes scientific outputs of all kinds citable, shareable and discoverable for the long term. Zenodo provides the following capabilities: sharing and linking research, citation and discoverability through the Digital Object Identifier (DOI) and harvestability via OAI-PMH by third parties, supports versioning, high reliability, trust and safety in data storage, article level, metrics, flexible licensing and supports FAIR principles (Findable, Accessible, Interoperable and Reusable).

2.2 Making data openly accessible:

- **Specify which data will be made openly available? If some data is kept closed provide rationale for doing so**
- **Specify how the data will be made available**
- **Specify what methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)?**
- **Specify where the data and associated metadata, documentation and code are deposited**
- **Specify how access will be provided in case there are any restrictions**

As a general note, FlexiGroBots data will be ultimately (after the pilot demonstrations have been run and finalized) offered for public access through the platform and open repositories such as Zenodo. This of course excludes private data (such as identities and contact details of application user data, etc.). All legal and other restrictions will be clearly outlined in the metadata.

Accessing data is most reliably realized by offering it through interfaces based on globally adopted standards. For geospatial or domain-specific data, this could be e.g. the OGC Web Feature Service for feature data, OGC Web Coverage Service for coverage data, OGC Web Map Service for maps, or the OGC Sensor Observation Service for sensor data. For most data access requirements, a standard already exists. Where additional requirements arise from the research in FlexiGroBots, the requirements shall be used to advance and mature existing standards, rather than re-inventing the wheel.

2.3 Making data interoperable:

- **Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability.**
- **Specify whether you will be using standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies?**

Interoperability aspects will be considered in the context of FlexiGroBots, aiming to enable the maximization of the value of the data provided by the project through the utilization of common systems for transmitting and/or exchanging environmental information.

In case the use of standard vocabulary for metadata description will not be possible, a mapping of more common ontologies might be provided from ad hoc specialists through specific technical contribution.

Data can be made available in many different formats implementing different information models. The heterogeneity of these models reduces the level of interoperability that can be achieved. In principle, the combination of a standardized data access interface, a standardized transport protocol, and a standardized data model ensure seamless integration of data across platforms, tools, domains, or communities. When the amount of data grows, mechanisms have to be explored to ensure interoperability while handling large volumes of data. We will need to review this element during the course of the project. For now, data interoperability is envisioned to be ensured through compliance with internationally adopted standards.

2.4 Increase data re-use (through clarifying licenses):

- **Specify how the data will be licenced to permit the widest reuse possible**
- **Specify when the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed**
- **Specify whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why**
- **Describe data quality assurance processes**
- **Specify the length of time for which the data will remain re-usable**

The reuse of data is a key component in FAIR. It ensures that data can be reused for purposes other than it was initially created for. This reuse improves the cost-balance of the initial data production and allows cross-fertilization across communities. FlexiGroBots will advertise all the data produced to

ensure that they are known to wider audience. In combination with standardized models and interfaces as described above and complemented with metadata and a catalogue system that allows proper discovery, FlexiGroBots can serve as valuable input outside of the project.

At this stage, it is not clear what licensing models need to be applied for the various data products produced in FlexiGroBots. Generally, the focus shall be on public domain attribution and open licenses that maximize reusability in other contexts. All data products produced by FlexiGroBots will be reviewed for FAIR principles once a year by the data producing organization. On the other hand, FlexiGroBots is open to any third-party data and process provisioning. Data quality is a key component for data reuse. Without proper quality parameters, data cannot be integrated into external processes, as the level of uncertainty of the remote processes becomes undefined. FlexiGroBots will review its data products for quality information provided as part of the metadata. Currently, ISO quality flags might be envisioned to be used.

3. Allocation of resources

Explain the allocation of resources, addressing the following issues:

- **Estimate the costs for making your data FAIR. Describe how you intend to cover these costs**
- **Clearly identify responsibilities for data management in your project**
- **Describe costs and potential value of long term preservation**

There are limited funds for open-access available in the FlexiGroBots budget since only several partners budgeted it upfront. Therefore, consortium partners will be either aiming towards enabling open access of peer-reviewed scientific publications through free options: deposition of the accepted manuscript in author’s personal web pages, institutional publication repositories and free public repositories while applying for green open access, or might use some amount from their other direct costs. In cases when journals chosen for optimal dissemination of scientific results (timely and wide dissemination leading to significant benefit to the community) are not offering green open access or the embargo period is too long, FlexiGroBots consortium might opt for the gold open access.

FlexiGroBots consortium is aware that costs related to data management in H2020 are eligible for reimbursement during the duration of the project. To ensure Open Access after the life time of the FlexiGroBots project, the FlexiGroBots consortium will whenever possible, use free tools that are compatible with the requirements by ORDP in H2020 (e.g. deposition of data and publications in OpenAIRE’s Zenodo centralized repository that allows for free and long-term deposition (for the next 20 years at least along with the experimental programme of its host lab CERN).

4. Data security

Address data recovery as well as secure storage and transfer of sensitive data

The FlexiGroBots project beneficiaries guarantee that all data collected during the project will be kept secure and unreachable by unauthorized persons. The data will be handled with appropriate

confidentiality and technical security, based on partners' best practices.

As an example, BioSense Institute will store all generated/collected data (both collected in the course of pilot 3 and as a part of Data Management activities) on a dedicated Data Storage System with dual controllers and a dual power supply. Everything stored on those machines is copied on at least three Hard Disc Drives (HDD). In case of failure of one of the HDD, data are secured on two others and within 24 hours the replacement HDD is obtained from the manufacturer. In the case of electricity cut-offs, dual power supply enables continuum by automatically swapping from electric network to UPS with diesel aggregate.

The data stored in the BioSense Institute Data Storage System are not exposed directly to the end-users/internet thanks to two-line defence architecture (Figure 1). In the first line, there is one Virtual Machine running as a Proxy server for all requests, also taking care of balance load. Calls are then forwarded to another Virtual Machine that can access the stored data. Thanks to such architecture, even if someone manages to intrude into the Proxy machine, it will not have direct access to the data, which are hidden behind another Virtual Machine.

The protection of data will also be ensured through procedures and appropriate technologies, like the use of HTTPS protocol for the encryption of all internet transactions and appropriate European and Internet security standards from ISO, ITU, W3C, IETF and ETSI. More specifically the server onto which the data will be stored will have server-side encryption allowing administration personnel to generate private keys for data access without access to the data themselves. That means that only authorized personnel will have access to the data and even in the case of a possible data leak or server hack the data stolen will be fully encrypted and thus non-accessible.

Transfer of (sensitive) data

Data transfer to and from end-users (including transfer of sensitive data if allowed) is performed encrypted, either sent by encrypted ZIP or RAR files or download directly as web-based services from servers (e.g. GeoServer). In any case, a strong password (more than 30 randomly generated characters in order to prevent dictionary or brute force attacks) is required for accessing transferred dataset and passwords must be sent separately from the dataset (preferably using also different channels of communication e.g. SMS, Viber, WhatsApp).

Prior the sharing for the analysis all data containing sensitive personal information has to be anonymized. Anonymization refers to removing any identifier that can reveal identity of the participants both from data and metadata.

Data protection

All the FlexiGroBots project related activities will be carried out ensuring the respect of all the ethical principles and in accordance with the **Directive 95/46/EC of the European Parliament** , about the protection of individuals with regard to the processing of personal data and on the free movement of such data, as well as the **Directive 2002/58/EC** concerning the processing of personal data and the protection of privacy in the electronic communications sector, as modified by the **Directive 2009/136/EC**. **All national data protection and privacy laws of Serbia will be equally followed and respected** (Official Gazette of RS no.97/2008, 104/2009, 68/2012 and 107/2012).

Moreover, the protection of personal data will also be ensured through **procedures and appropriate technologies,** like the use of HTTPS protocol for the encryption of all internet transactions and appropriate European and Internet security standards from ISO, ITU, W3C, IETF and ETSI.

Based on the FlexiGroBots Grant Agreement and corresponding activities, **no collection or processing of personal sensitive data** (e.g. health, sexual lifestyle, ethnicity, political opinion, religious or philosophical conviction) **is or will be involved.**

Data retention

To ensure privacy, **all data will be anonymized, encrypted and stored on a server to which only the relevant staff have access**. In particular, the server in which the data will be stored will have **server-side encryption.** That means that the server's administration personnel will be able to generate public keys for specific personnel who will have access to the data but will not be able to

access the data themselves (since the private keys required for this access will be generated on the machine of the person with access to the data). Therefore, **only the required personnel will have access to the data and even in the remote case of a possible data leak or server hack the data stolen will be fully encrypted, and, thus, fully non-accessible.**

Data destruction

All individuals **will have the right to obtain the erasure of personal data relating to them** and **the abstention from further dissemination of such data** according to the GDPR*; Article 17* ***Right to be forgotten and to erasure****).* **They will be informed in advance about this right in the information sheets**. Applications for the erasure of data will be carried out without delay. **In case the personal data have been made public, the FlexiGroBots consortium will take all reasonable and necessary steps, including technical measures, to inform third parties, which are processing such data that a data subject requests them to erase any links to, or copy, or replication of that personal data.** A procedure for exercising ***the right to be forgotten and to erasure*** will be provided, and will include, checking the validity of the request, identifying data which should be erased, monitoring the erasure process, and informing the person in question.

In case of conducting research on human participants (on any kind), **FlexiGroBots consortium confirms in writing that it will be compatible with EU and international laws, Horizon 2020 ethical standards, as well as national legislations of the Republic of Serbia** and that **it could have been legally conducted in** (at least) **one of the EU Member States.**

Curation, preservation and dissemination

As an example of BIOS as responsible for data management plan and execution, all data collected/generated by its employees and during its activities are stored at the institutional servers. The data are kept at least for 10 years after production and could be made accessible to third parties upon request and after approval from the Institute Scientific Board.

It is advised to all the partners to use secure and trustworthy tools for the data to curate and preserve datasets collected/generated during the lifetime of the FlexiGroBots project, such as B2SHARE (https://b2share.eudat.eu/) or similar. Apart from facilitating storage and long-term persistence of data, B2SHARE also allows worldwide sharing within scientific and citizen scientist communities. It is a free service that allows users owners to define access policy and it includes a built-in licence wizard that will facilitate the selection of an adequate licence for research data. Another advantage of B2SHARE usage is automatic assignment of PID (persistent identifier) and DOI codes for easier referencing in scientific papers.

Data underlying peer-reviewed scientific publications are curated, preserved and disseminated through centralized repositories and if possible, as supplementary material to the published paper.

5. Ethical aspects

To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former

FlexiGroBots' goal is to make heterogeneous multi-robot systems cost-effective and to deliver more flexible solutions by i) employing multiple heterogeneous versatile (e.g. multi-task) robots, ii) which collaborate to accomplish complex missions; iii) ensuring scalable human oversight and intervention through adaptive mission control mechanisms (i.e. without information overload/overwhelming effort for the farm operator); and iv) enabling an agricultural data economy to profit directly—by offering and selling data gathered in the field—and indirectly, by acquiring valuable datasets and/or machine learning models—already trained from them—to build artificial intelligence capabilities for the robots within the system. The main research objectives are related to robotics and the project does not include vulnerable populations in the study. In addition to the technical research, FlexiGroBots will also

analyse each pilot according to the EU guidelines for trustworthy AI (especially the operationalized assessment list). We will take into account issues of human agency, technical safety, privacy, transparency, diversity/non-discrimination/fairness, societal and environmental wellbeing and accountability. We will develop recommendations and share our ethics research beyond the consortium. We aim to generate new knowledge on the ethical implications of robotics and AI and this research is an important part of each pilot (see T4.4, T5.4, T6.4) and also the final results of the project (see D7.7 "Report on ethical AI use in agriculture robotics").

As part of these tasks, FlexiGroBots will take into account the work done within the recently completed REELER (Responsible Ethical Learning with Robots) project, which will help to deal with broad ethical issues and especially for the ratio of human-robot proximity and its impact, which are often unforeseen. REELER recommendations cover standard ethical concerns and they will be enhanced through institutional and national levels of clearance.

Human participants

FlexiGroBots foresees a set of activities which involve human participants in its pilots.

One of the tasks that will be carried out in **Pilot 1** is to assist the operators in the manual harvesting of grapes, working with small robots to assist the operators in baskets transport (human-robot interaction). **Pilot 2** will monitor people on the field due to safety reasons by cameras on UAVs, at a distance from which it would not be possible to identify them personally.

The tasks within these pilots will be conducted under ethical considerations and following the approval by the research ethics committees including plan to ask consent from all the workers moving at the pilot sites while pilot operations run for their permission to be recognised as humans and not as persons, and also for making everyone aware of potential risks related to autonomous machines moving on the field.

FlexiGroBots project will incorporate engagement of end-users, on-field demonstration and gathering user experiences and attitudes towards technologies being developed (conducting surveys, interviews or focus groups where personal information will be gathered and stored). During the pilot’s assessments interviews with the technical consortium partners are planned. The FlexiGroBots team will have to pay attention to privacy, data protection, data management and the health and safety of participants. No issues related to health and safety have been detected and the questions related to privacy and data protection and data management are described under “Personal data” in the next section.

During the trials in pilots and the research related to this project, FlexiGroBots will ensure:

- respect for people;
- respect for human dignity;
- fair distribution of the benefits and burden of research. The general principle is “maximise benefits and minimise risks/harm”;
- protection of the values, rights, and interests of the research participants: the research methodologies will never result in discriminatory practices or unfair treatment.

The participation of human subjects in any study must be voluntary, based on informed consent, governed by review and approval by a local Institutional Review Board.

Personal data

In those cases where personal data are involved (previously indicated by each partner in the data summary) detailed information is planned to be provided:

1. on what personal data are collected, stored and processed;
2. on the recruitment process, inclusion/exclusion criteria for participation;
3. on privacy/confidentiality and the procedures that are implemented for data collection, storage, access, sharing policies, protection, retention and destruction during and after the project;
4. on how informed consent is pursued;

5. if application/is needed to be filed with a local/institutional ethics review body (if personal data is being collected) and if yes, which bodies/where/when.

The legal experts of those partners that process personal data will guarantee that this process, including the information for the individuals about data protection issues, fully complies with national and EU laws.

There are several situations related to the management of personal data in FlexiGroBots:

1. Like any other research project, FlexiGroBots involves, in the first place, processing of personal data relating to conducting the action itself: identification data of consortium members and researchers, data concerning advisory board members, etc. In this case, the Data controllers will be the consortium partners as far as they process data related to their own staff members, the project coordinator as far as it processes data in the context of project management tasks and the European Commission as controller of the Horizon 2020 participant portal. Although this kind of processing personal data is self-evident and unavoidable in the context of an H2020 action, it falls under the scope of application of the European and national data protection legislation and, thus, requires e.g. full respect of all relevant provisions of the GDPR.

2. In the second place, FlexiGroBots will be confronted with personal data protection rules in the context of the pilot scenarios. The pilot validation will be based on datasets that may include personal data. Data controllers here are the owners of the datasets. Data processors are those project partners who don’t own the datasets but make use of them for the development of the project results. As of today, the following table summarizes the roles of partners involved in pilots:

Pilot	**Partner**	**Data Controller**	**Data Processor**
3	BIOS	x	x
2	VTT	x	x
2	MTE	x	x
1	SERESCO		x
1	TERRAS	x	
2	LUKE	x	x
2	PROBOT	x	x
1	CSIC	x	x
1	WUR	x	x
3	ART21	x	x
1-3	ATOS		x

Table 2: Roles of partners within pilots

Here we note that there is a possibility of changes during the lifecycle of the project. In case of pilot 3, that includes non-EU and EU country, Serbia and Lithuania each party will comply with all applicable laws and further data transferred will be in anonymous form.

3. The pilot validation may require the involvement of external participants in the project. This participation may require the collection of some personal data related to the recruitment, even when this dataset does not apply for the validation of the pilot itself. This includes data like name, affiliation and email address.

For the normal implementation of the action related to project coordination, internal communication, and project communication and dissemination, specific consent will be requested to project members and project supporters (e.g. Advisory Board members) in accordance with a privacy policy for the use of personal data. This privacy policy and the consent will include aspects such as the use of data, access to data, retention date and users’ rights regarding their data.

The project website will also provide a privacy policy describing data protection regarding the services

offered to web users through the website (e.g. web account, newsletter subscription).

For the implementation of the technical activities (e.g. development and validation), any shared information that is made available between consortium partners (and their third parties involved in the action), like background, results, confidential information, datasets, or any data or information, shall not include personal data. Each partner will take all necessary steps to ensure that all personal data is removed, obfuscated, or made inaccessible from the shared information, prior to providing it to any other partner. All the aspects related to this issue will be considered and regulated by the FlexiGroBots Consortium Agreement.

To make the data inaccessible, partners will follow the guidance set forth in the Article 29 Working Group 05/2014 Opinion on Anonymization Techniques, and specifically will explore their recommendations on Pseudoanonymization, Noise addition, Substitution, Aggregation, K-anonymity, L-diversity, Differential privacy and Hashing/Tokenization, even homomorphic encryption and multiparty computation techniques if needed.

Further information on personal data

The obligations of the partners processing personal data are regulated in the General Data Protection Regulation - GDPR [5]. These obligations involve both EU members and non-EU members, and are also described by the Ethics and Data Protection Guidance edited by the European Commission (version 14/11/2018) [6], which is not a legal official document but describes in a clear, accurate and specific manner the requirements and obligations to be accomplished under the H2020 framework, while processing personal data under the GDPR.

Following the requirements of the GDPR and the Ethics and Data Protection Guidance, and although this matter is contemplated and regulated in Annex I Part B of the Grant Agreement (Section 5. Ethics and security) – without limiting, restricting or contradicting what it is written in them – in this section we briefly describe the considerations to be taken into account for all the partners that process personal data, even if they have access to them, and at which compliance and observance are committed.

Whenever possible, and as it is established in Annex 1 Part B of the Grant Agreement, the personal data of the participants in the research will be anonymized, in the way that data will not relate to identifiable persons, at the moment and at the time that the data is collected, remaining in this way out of the GDPR application; in any other case, all partners will be subjected to the obligations established in the GDPR in the matter of processing personal data of the participants in the research, and in this way:

1. The partners must treat the data in accordance with the obligations that are established in the GDPR, and in accordance with the principle of minimization:

The partners must apply the GDPR when processing personal data that they have collected or accessed (understood as personal data any data that can identify the identity of a physical person, according to article 4.1 of GDPR and understanding by "treatment", according to article 4.2 GDPR, the collection, registration, organization, storage, structuring, adaptation or alteration, recovery, consultation, use, disclosure by transmission, dissemination or availability, alignment or combination, restriction, deletion or destruction of personal data).

Under the principle of minimization and proportionality of data, the treatment must include only the data necessary and proportional to achieve the purpose for which they were collected, that is, to meet the objectives of the research, (article 5 GDPR), and for the duration of the project -or for the time necessary to fulfil other obligations established for it (ex, audits). And if the purpose for the processing cannot be fully identified at the time of data collection or if it is necessary to keep them beyond the duration of the project, the reasons for such conservation and data collection for longer than the project must be explained.

Likewise, the principle of minimization applies to the entire data processing process (not only to the amount and adequacy of data collected for the established purpose) as well as the design of the procedures for their access, transfer or communication, when sharing data with other partners, storage and when establishing the conservation period. Most appropriate technical and organizational

measures must be applied in each case to comply with the principle in order to ensure and protect the fundamental rights in case of unauthorized access.

1. The partners must inform the participants and obtain their specific consent to collect and treat their data:

The partners must inform the participants of the investigation of the destination, use, purpose, treatment, conservation, and possible assignments or international transfers (in case) on the collected personal data (article 13-ss GDPR) and obtain from them their consent to be included in the research project, as a previous priority to the processing of the data, and as a legal or lawful basis for the treatment of them (article 6 GDPR).

Only data will be collected with a strict connection with the objective of the project. In case the data will be used for another purpose(s), in multiple research projects or in future research projects, if they will be shared with another project partner(s), or if they are going to be transferred to partners outside the EU, participants in the research must be informed before providing their consent for the processing of their data (article 13 GDPR). It is recommended to request additional and explicit consent for this secondary use of the data and give participants the opportunity to choose not to participate in the additional processing operations of their data.

In case that it is not possible to anonymize personal data, partners should pseudonymize them.

If it is not possible to anonymize the personal data of the participants in the investigation at the moment of collecting the data, in case that it is necessary to maintain a link between the research subjects or participants and their personal data, these data should be pseudonymized in order to protect their fundamental rights in case of unauthorized access to them, using techniques such as coding or hash.

1. Partners must implement the most adequate technical and organizational measures to protect the personal data to be treated.

Partners must implement, describe, or detail the technical and organizational measures most suitable to protect the personal data that they process, and avoid their disclosure, unauthorized access to data and / or equipment, accidental elimination or destruction.

To do this, they can use the following measures (listed without limitation): pseudonymization or anonymization of data, applied cryptography of data, and devices in which they are stored (encryption and hashing), the establishment of procedures to limit the use of data, ensure that all partners, contractors or service providers that process research data comply with the GDPR, data minimization (in the terms set out above) and adequate protection of passwords and passwords, or limit access to certain people.

In case that a partner treats large-scale data (24 and 91 GDPR), it must specify/explain how they will guarantee an (improved) level of data security; and if techniques such as data mining, “web crawling” or social media analysis are used in large-scale data processing, or the project involves intensive monitoring or follow-up of research participants (for example, regarding the behaviour or activities) you must specify/ explain how you will take measures to protect your personal data and your fundamental rights. and in case that the research project involves the automated processing or the elaboration of personal data profiles, the interested parties should also be provided with information about the data processing, the elaboration of automated profiles and the evaluation.

1. Partners must keep personal data only for the strictly necessary duration:

Collected personal data must be stored only for the necessary time for the purposes for which they were collected, or according to the provisions concerning audits, storing, or retentions established for the project. they should be deleted or be destroyed safely and unrecoverable when there are no longer needed. while they are retained or preserved for the audit processes, they should be treated only for that purpose and stored safely.

If personal data have been stored in the cloud or by an external service provider, shared with other partners or transferred to a third party during the project, it must be ensured that they have been

removed (unless there is a legitimate basis to retain them). It is advisable to demand reliable evidence of the destruction by means of a responsible declaration issued for this purpose.

1. In case of sharing data with another(s) partner(s)

In principle, the personal data collected and processed by the partners may not be transferred to third parties, except if they are shared or communicated to any other Project partner, if applicable, in order to carry out their tasks within the Project. In this case, if a partner shares with other partner(s) of the consortium, both have the responsibility to process the personal data collected in the course of the Project (they could have joint data controllers) and the responsibilities of each partner must be collected or established in an agreement available to the partners.

1. In the case of international transfers.

There is international data transfer in the case that partners or service providers outside the EU can access personal data collected in the EU. This includes the case that non-EU partners access a platform containing personal data. In this case, the GDPR must be accomplished.

The H2020 standards and ethical guidelines apply to all partners, regardless of the country in which the research is conducted: thus, the GDPR applies to all data controllers, public or private, belonging to the EU or non-EU, that treat and maintain the personal data of the interested parties residing in the EU. Therefore, partners or service providers from non-EU countries must comply with the GDPR.

If the personal data of research subjects are compiled in non-EU countries, the requirements established by the EU must also be applied, so the GDPR applies to all personal data processing operations performed by data controllers with EU headquarters, regardless of where the data processing takes place.

National data protection laws of the country in which the investigation is performed must be complied with (for example, notify or request permit or additional authorizations to the authorities in the matter...). The partners or service providers of non-EU countries must comply with the GDPR and, where appropriate, with their national standards in the field of personal data protection.

1. In case that some partner use data collected previously

Only data with a strict connection with the objective of the project will be collected.

In point V of the Guide ethics and data protection [6] it is addressed how to proceed in the following cases (reading is recommended):

The use of data collected for a purpose and then used for other research processes without the consent or knowledge of the interested party:

    1. The use of data that are publicly available
    2. The use of the data that the interested party has made public
    3. The use of open-source data on identifiable persons, creating new records or files / profiles (use of social media data) without requesting consent from the affected / interested party for the use of their data
    4. The use of personal data that was collected from a previous Research Project
    5. The use of personal data provided by a third party and those interested have not given prior consent to the use in the investigation or treatment

1. Obligation of compliance by partners of the European and national legislation in the matter of protection of personal data and responsibility of partners regarding such compliance.

The exposed obligations do not prevent or substitute the application of the other obligations established by the GDPR or that must be applied in compliance with the regulations in the matter that governs the countries that do not belong to the EU (Serbia), and neither exempt nor limit the responsibility of each partner in the treatment and protection of personal data obtained (by direct collection or access to them through the transfer or communication, or by sharing them with another partner), in accordance with the European and National regulations governing this matter. Therefore,

each partner is responsible for the data processing that it carries out.

Non-EU countries

The H2020 ethical rules and guidelines apply to all partners, regardless of the country in which the research is conducted: thus, the GDPR applies to all data controllers, public or private, belonging to the EU or non-EU, that process and maintain the personal data of the interested parties residing in the EU. Therefore, partners or service providers from non-EU countries must comply with the GDPR. And if the personal data of research subjects are collected in non-EU countries, the requirements established by the EU must also be applied, so the GDPR applies to all personal data processing operations performed by data controllers with EU headquarters, regardless of where the data processing takes place.

Serbia

All activities related to data protection will be performed in accordance with the Serbian Law on personal data protection (Official Gazette of RS no.97/2008, 104/2009, 68/2012 and 107/2012), which in Article 10 emphasizes that written consent to data processing is deemed to be valid if given by a person who has received prior information from the collector of the data. Article 15 of the same Law provides the details on what this prior information has to include (e.g., the identity of the interviewer, purpose of data collection/processing, how data will be used, who will use the data, is data provided on voluntary base, etc.).

All the work that will be conducted in Serbia will follow the procedures and criteria that have been set and are in accordance with standards and guidelines of H2020 program, EU legislation, national legislation in Serbia, professional standards and law of the Republic of Serbia (and Statute of the organization in case of BIOS).

Information sheet and informed consent

Before signing the informed consent form, potential participants are informed about the study by an information sheet, which is a layman explanation of the research. The information provided to the subjects is intended to give each participant thorough understanding of the purpose, nature, methods, procedures, possible risks and benefits of the examinations, as well as the planned use of the data to be collected to make an informed decision as to whether to participate in your research project. It will also provide potential participants with details of sources of further information to answer any further questions that they might have.

These sheets will include information such as the following:

- Details of the research project and its main objectives.
- The purpose of the research.
- What participation will involve.
- The benefits and disadvantages/risks of participation.
- A clear statement that participation is entirely voluntary and that participants can withdraw from the project at any time without prejudice, now or in future.
- Details of what will happen to the data collected and the results of the research.
- Details of who to contact for further information or for additional clarifications.

The participant will be given one copy of the signed original information and consent form; the original will be kept by the investigator.

The **informed consent forms** and detailed **information sheets in FlexiGroBots**:

- are written in clear language and in terms that the subjects will fully and easily understand and will avoid long illegible terms and conditions full of legalese. It will be distinguishable from other matters and must be as easy to withdraw consent as it is to give it;

- describe the aims, methods and implications of the research, the nature of the participation and any benefits, risks or discomfort that might have because of participation, without the necessity of giving a reason, without further consequences for them;
- explicitly state that participation is voluntary and that anyone has the right to refuse to participate and to withdraw their participation, samples or data at any time — without any consequences;
- state how samples/data will be collected, protected during the project and either destroyed or reused subsequently;
- state what procedures will be implemented in the event of unexpected or incidental findings (in particular, whether the participants have the right to know, or not to know, about any such findings);
- will be signed in writing by participants.

Informed consent

The consent form should be a short document that concisely covers the core statements to which the participant is being asked to agree in clear and concise language.

- **FOREWORD -** to understand the purpose, context and scope and context
- **PURPOSE OF THE PROJECT -** to briefly describe the purpose
- **PURPOSE OF THE ACTIVITY-** to briefly describe purpose of the activity
- **EXERCISE PROCEDURES –** to list all procedures, preferably in chronological order, which will be employed in the activity. If audio taping or videotaping are going to be used, information about the use of these products will be provided
- **RISKS -** to list all reasonably foreseeable risks, if any, of each of the procedures to be used in the activity, and any measures that will be used to minimize the risks
- **BENEFITS –** to list the benefits you anticipate will be achieved from this validation, including benefits to participants, others, or the body of knowledge.
- **CONFIDENTIALITY –** to state measures taken to ensure confidentiality
- **WHAT WE WILL DO WITH YOUR DATA** – to provide information on data storage, anonymization, sharing, defined period after data will be destroyed after
- **CONTACT INFORMATION -** to provide contact information for questions regarding participation in the activity
- **VOLUNTARY PARTICIPATION** - to highlight that participation is voluntary with freedom to withdraw at any time and without giving a reason

A template for the Inform consent can be found in Section 7.3.

6. Other

Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any)

None