D02: DATA MANAGMENT PLAN (DMP)

Date: 30.10.2020.

Name of the Program: PROMIS
Name of the Project: **Prediction of Cancer Treatment Effectiveness with Stimuli-responsive Nanomaterials**
Project Acronym: **PRECAST**
PI **Nikola Knežević**
Scientific Research Organisation (Lead SRO) **Biosense Institute**
Lead SRO Acronym BIOS
Implementation period: **01/09/2020 – 31/08/2022**

List of figures

List of tables

1. Introduction

1.1. Purpose of the document

The PRECAST project is funded by the Science Fund of the Republic of Serbia -PROMIS program, aiming to develop methodology for predicting the effectiveness of targeted cancer therapy by means of magnetic resonance imaging (MRI) and application of **cancer-targeting nanomaterial containing loaded contrast agents (CAs)**. This would enable strong contrast enhancement selectively at the tumor tissues in case of successful cancer-targeting, which would reveal the most potent composition of a drug-filled, structurally analogous, nanomedicine for yielding the most potent cancer targeting and **personalized treatment of cancer**. Partner on the Project is Faculty of technology, University of Belgrade, while supporting foreign institutions are Trinity College Dublin and Wageningen University.

This deliverable constitutes the Data Management Plan (DMP) of the PRECAST project. It is the first version of DMP, describing data management life cycle for all research and publication data. **This deliverable is introducing ways to produce, collect, process, save, exchange, share, and publish research data**.

Other aim of this deliverable is to set communication and establish data handling, sharing, storing and usage between project partners and to create open, clear, and understood procedures that will be followed by the entire PRECAST consortium.

The PRECAST coordinator will update this DMP, if there is a need for improvement considering changes and activities during the project's lifespan. For example, new ways of data generating or collecting; changes in data storing and saving methods; changes in consortium policies regarding open access and background data; if specific part of the openly accessible research data can jeopardize main objectives of the Project research; other factors that will make significant changes in project activities, timeline, reporting and result sharing.

This DMP will be evaluated several times to ensure the smooth delivery of the goals highlighted in this document.

1.2. Definitions

Open Access refers to the practice of providing online access to scientific information that is free of charge to the end-users and reusable. Scientific information, in the context of research and innovation, can mean:

- Research data refers to the information, in particular facts or numbers, collected to be examined and considered as a basis for reasoning, discussion, or calculation (i.e. statistics, results of experiments, measurements, survey results, interview recordings and images). Here, the focus is on research data that is available in digital form and users can access, reproduce, and disseminate openly accessible research data free of charge.

- "Peer-reviewed publications" refer to those accessed by other scholars. It is focused on journal articles, but not exclusively. It can also include other types of scientific publications such as:

monographs, books, conference proceedings and gray literature (informally published written material not controlled by scientific publisher, e.g. reports).

2. PRECAST data summary

2.1. Purpose of the data collection

PRECAST is a research project aiming at developing methodology for predicting the effectiveness of targeted cancer therapy by means of magnetic resonance imaging (MRI) and application of **cancer-targeting nanomaterial containing loaded contrast agents (CAs)**. This would enable strong contrast enhancement selectively at the tumor tissues in case of successful cancer-targeting, which would reveal the most potent composition of a drug-filled, structurally analogous, nanomedicine for yielding the most potent cancer targeting and **personalized treatment of cancer**.

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, will be integrated into the PRECAST platform. During the lifecycle of the PRECAST 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.

2.2. Types, sources, formats, and estimated size

Many of the activities of PRECAST projects will produce research data. At this early stage of the project implementation, it is difficult to anticipate all types, formats and size of data PRECAST project team will be dealing with. This will largely depend on the joint research strategy, acquired technology (i.e. capital equipment and infrastructure), researchers' competences and newly established collaborations (i.e. research projects, industry, governmental bodies, NGOs) which all will be known later in the course of the project implementation.

Details on the data that is expected to be collected and generated by PRECAST team during the implementation of the project are summarized in the Table 1.

Type	Formats	Estimated Size (per dataset)
Data from laboratory experiments (e.g. MRI measurements, SEM images, FTIR, TGA, UV/VIS, fluorescence, TEM, XRD analysis)	.txt, .csv, .tiff, .jpg, .doc, .xls, .s1p, .s2p, .raw, .rd5, .h3d	1GB
Publications (e.g. reports, deliverables, dissemination material)	.doc, .pdf, .ppt, .ai, .psd, .mp4	10MB (100MB for videos)

Table 1: Types of collected/generated datasets, their formats and estimated size.

2.3. Roles and responsibilities

With respect to the PRECAST, BIOS, as the Coordinating Legal Entity assumes the main role and responsibilities of a Data Controller, through the appointed Data protection officer (DPO). In the course of PRECAST, the owners of the datasets are treated as Data controllers, while data processors are those project partners who don't own the datasets but make use of them for the development of the project results (see Table 2 for more details). All the Consortium partners involved in data processing will also maintain records as data processors.

Each partner is responsible for the data generated from his/her institution as data producers, proper data collection, documentation, and storage throughout the duration of the project, under the supervision of their respective DPOs. In addition, each partner is responsible for informing their own staff involved in the PRECAST project about the need to comply with the legal principles and provisions with regard to data processing.

The party responsible for the file and, if applicable, the party in charge of the processing must implement technical and organizational measures necessary to guarantee the security of the personal data and prevent their alteration, loss, unauthorized processing or unauthorized access, taking into account the state of technology, the nature of the data stored and the risks to which the data are exposed, whether due to human action or to the physical or natural environment.

Specific agreements may be signed among partners in order to grant access to the different datasets for different uses (data storage, data processing, service provision).

3. FAIR data in PRECAST project

3.1. Making data findable

3.1.1. Naming conventions

Following remarks regarding the name of electronic records (files) should be followed:

- use _ instead of space
- preferably not exceed 255 characters (to ensure it is readable at 32bit and above operating systems)
- if the document is modified, it contains version number and the date of the last modification
- contains all denominators required for identification of the file content.

E.g. For SEM images the file name should contain the name of the sample, the name of the researcher, magnification, location taken, date of the record (e.g. YYYY_MM_DD). If some of the above numbered information is not adequate for description of content it should be adjusted (e.g. in case of other types of measurements).

For deliverables: deliverable number and official name as in the Grant Agreement should be part of the file name. The author of the file puts the initials of his/her name in the file name. Each file name contains the initials of the name of its last author (e.g. D1.2 Quality_Management_Plan_v1_MM_20.10.20.doc - The

deliverable 1.2, during elaboration of v2, last author Mickey Mouse, date of last modification: 20.10.20, before v2 being released by Project Coordinator)

3.1.2. Metadata

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

Collected/generated and re-used data can have associated metadata directly or indirectly. Directly means that certain file types support metadata natively (e.g. JPEG files support metadata that include information such as image size, creation tool, compression ratio, original source), while indirect metadata are created by the producer using tools by which source files are generated or used. For generated data and publications (both peer-reviewed and internally reviewed), metadata elements in accordance with the guidelines of the repository will be included. If such guidelines do not exist, commonly accepted metadata for web resources should be used. If this is not possible, creation of more suitable dataset will be taken into consideration.

Apart from metadata required by the repository, publications must be accompanied also by the bibliographic metadata in a standard format including all of the following: the terms, the name of the action, the publication date, and length of embargo period if applicable, and a persistent identifier.

All peer-reviewed publications will be identifiable and locatable by means of persistent and unique Digital Object Identifier (DOI)[1], assigned by the publisher. In addition, a list of keywords will accompany the publications to optimize search in the domain of interest and thus expand possibilities for re-use.

For research data, wherever possible/applicable, following metadata standards will be encouraged:

- CERIF - Common European Research Information Format. CERIF is the standard that the EU recommends to its member states for recording information about research activity.
- CIF - Crystallographic Information Framework. An extensible standard file format and set of protocols for the exchange of crystallographic and related structured data.
- Observations and Measurements. This standard specifies an XML implementation for the OGC and ISO Observations and Measurements (O&M) conceptual model, including a schema for Sampling Features.
- CSMD-CCLRC Core Scientific Metadata Model. A study-data oriented model that captures high-level information about scientific studies and the data that they produce, primarily tailored for the physical sciences.
- MIBBI - Minimum Information for Biological and Biomedical Investigations. A common portal to a group of checklists of Minimum Information in nearly 40 biological disciplines. DDI Metadata - Data Documentation Initiative (International standard for describing the data produced by surveys in social, behavioural and economic sciences).
- OME-XML - Open Microscopy Environment XML. A metadata standard and data file format for biological light microscopy data.

[1] A Digital Object Identifier (DOI) is a serial code used to uniquely identify objects. Further information can be found at https://en.wikipedia.org/wiki/Digital_object_identifier.

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 a shared drive, 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 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).

3.2. Making data openly accessible

Open Access of data goes in line with addressing management of knowledge and Intellectual Property Rights (IPR) as well as deciding on mechanisms for knowledge sharing and exploitation.

Openness is the route to scientific excellence, as it fosters knowledge circulation, allows building on previous research results, and encourages collaboration. Project partners fully endorse Open Access to their scientific results and research data. All beneficiary institutions have been encouraging authors to publish their research according to Open Access principle and thus this DMP is in line with procedures regarding Open Access of scientific data established in PRECAST project partner institutions.

BioSense Institute (BIOS) dissemination efforts and its focus on collaborations with international partners, together with the dedication to design our research agendas in alignment with the end-user needs, ensure a constant bidirectional flow of knowledge. BIOS researchers publish their work in scientific journals with large audiences to increase the impact of their work and in order to make it visible we have an established practice of green open access to scientific publications via the Institute's web site, but also via repositories available through OpenAIRE and professional social networks (i.e. ResearchGate, LinkedIn, Academia). By participation in H2020 eLTER and H2020 EUDAT2020, BIOS is not only familiar with Pilot on Open Research Data, but it is also active in providing new approaches to data management and multiple widespread use of a distributed ecosystem research infrastructure and data it collects. Through H2020 EUDAT2020, BIOS is involved in a joint strategic effort to enable European researchers to preserve, find, access, and process data in a trusted environment, as part of a Collaborative Data Infrastructure.

3.2.1. Data open Access

Data taken from open sources, so as all data collected/generated by PRECAST project that underline peer-reviewed scientific publications, will be made openly available as the default. As can be seen in Table 1 all

data are in common formats which are compatible with various software tools. In most cases commonly used licensed software has its open source alternative which widens and simplifies access to collected/generated data. In case specific tools are required to explore and manipulate collected/generated datasets, whenever possible access to the relevant software (e.g. in open source code) will be provided along the dataset.

3.2.2. Publications Open access (including underlying dataset)

We are aiming for visibility of our research and its impact in respective fields of science, therefore we identified an indicative list of target scientific journals for disseminating the scientific output of PRECAST project (PRECAST Grant Agreement, Measures to maximise impact). Publishers have different policies on self-archiving and open access (Table 2) and therefore researchers are requested to have this in mind before copyrights are transferred to the publisher. This will allow for choosing the best option for timely archiving in repositories (including public repository Zenodo) and open access of their publication (including underlying data).

Publisher	**Embargo for open access from public repository**	**Reference**
American Chemical Society	No	http://pubs.acs.org/page/copyright/journals/faqs.html
American Physical Society	No	https://journals.aps.org/edannounce/PhysRevLett.106.070001
BioMed Central	No	https://www.biomedcentral.com/about/open-access
Elsevier	0-36 months	https://www.elsevier.com/__data/assets/pdf_file/0005/78476/external-embargo-list.pdf
Emerald Group Publishing	0-24	http://www.emeraldgrouppublishing.com/openaccess.htm
Hindawi*	No	https://www.hindawi.com/license/
IEEE	No	http://www.ieee.org/publications_standards/publications/rights/rights_policies.html
MDPI*	No	http://www.mdpi.com/authors
Nature	6 months	http://www.nature.com/authors/policies/license.html
Optical Society of America	No	https://www.osapublishing.org/submit/review/copyright_permissions.cfm
PlosOne	No	http://journals.plos.org/plosone/s/licenses-and-copyright
Science	6 months	http://www.sciencemag.org/authors/science-editorial-policies
SPIE	No	https://spie.org/Documents/Publications/JournalsCopyrightTransfer.pdf
Springer	12 months	https://www.springer.com/gp/open-access/authorsrights/self-archiving-policy/2124
Taylor & Francis	12-18 months	http://authorservices.taylorandfrancis.com/sharing-yourwork/
Wiley-Blackwell	12-24 months	https://authorservices.wiley.com/author-resources/Journal-Authors/licensing-open-access/open-access/self-archiving.html

Table 2: Publisher's self-archiving policy of accepted versions of manuscript in the most important high impact scientific journals and magazines

PRECAST researchers will take the following journals as first choices for consideration for publishing – Table 3.

Journal	**Publisher**
Nature Materials	Nature
Nature Nanotechnology	Nature
Advanced Materials	Wiley
Advanced Functional Materials	Wiley
Journal of the American Chemical Society	American Chemical Society
ACS Nano	American Chemical Society
Angewandte Chemie (International Edition)	Wiley
Small	Wiley
Journal of Materials Chemistry. A	Royal society of chemistry
Nano letters	American Chemical Society
Journal of Controlled Release	Elsevier
Chemistry of Materials	American Chemical Society
ACS Applied Materials and Interfaces	American Chemical Society
Nanoscale	Royal Society of Chemistry
Lab-on-Chip	Royal Society of Chemistry
Journal of Materials Chemistry. B	Royal Society of Chemistry
Sensors and Actuators B	Elsevier
Sensors and Actuators A	Elsevier
Biosensors and bioelectronics	Elsevier
IEEE Sensors Journal	IEEE
ACS Sensors	American Chemical Society
Sensors	MDPI
Pharmaceutics	MDPI

Table 3: High impact scientific journals and magazines of interest for PRECAST

After decision to publish their results and choosing target journal for publication (focusing on the visibility and impact it will provide) the authors are advised to perform the following checks that will ensure the affiliated institutions are complying with legal requirements on open access to scientific publications:

Before submission: Explore publisher policy on self-archiving and Open Access. Please note that you are required to provide Open Access to your data at most 6 month after publication (12 months for publications in the social sciences and humanities). If only Gold Open Access is available, ensure that resources are allocated for paying article processing charges prior to submission.

After acceptance: Explore publishing agreement that is to be signed with the publisher.

Authors are encouraged to retain their copyright and grant adequate licenses to publishers. If needed request for amendment of publishing agreement to allow retaining (EC offers a model amendment at: http://ec.europa.eu/research/participants/data/ref/h2020/other/hi/oapilot/h2020-oa-guide-model-for-publishing-a_en.pdf).

Open Access for publications with internal reviewing process (e.g. presentations published in workshops, seminars and other related activities without a peer-reviewed process) could be uploaded directly to the publisher repository (e.g. repository with the conference proceedings), while the PRECAST project deliverables under the public status are uploaded to the precast website. PRECAST project dissemination material (i.e. press releases, brochures and other promotional material, videos) is stored into the PRECAST website.

As a general note, PRECAST 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.

For most data access requirements, a standard already exists. Where additional requirements arise from the research in PRECAST, the requirements shall be used to advance and mature existing standards, rather than re-inventing the wheel.

3.3. Making data interoperable

3.3.1. Compliance with standards

Due to the versatility of the activities, a number of standards will be used to ensure accessibility and reusability of generated data, such as ISO, W3C, OASIS, OGC, IEEE. In addition, common Internet security standards (e.g. ITU, W3C, IETF and ETSI) will be used.

All generated data are in standard formats for the given type of datasets making them usable by available software applications and re combined with datasets from different origins (e.g. SOS – Sensor observation service[2] to store environmental time series data to BioSense Institute server which makes it interoperable with any SOS compatible software). Common data formats ensure there are both licensed software applications and alternative open-source software for handling and analysis of collected/generated data.

3.3.2. Vocabularies

To allow for interoperability of data (both interdisciplinary and intradisciplinary), BioSense Institute will use standard vocabularies/ontologies for all data in collected/generated datasets.[3]

At this point we do not plan to generate uncommon, project specific, vocabularies/ontologies. In case this gets unavoidable we will create and provide mappings to more commonly used ontologies.

Interoperability aspects will be considered in the context of PRECAST, 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 contributions.

[2] *https://en.wikipedia.org/wiki/Sensor_Observation_Service*
[3] *https://admin.iprp.global/sites/default/files/2019-02/NWG_JRC-TechnicalReport_2018_2019_0205_0.pdf*

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.

3.4. Data re-use

3.4.1. Licensing for widening re-use

Before the data are shared for re-use, appropriate Intellectual Property Licensing is implemented. The databases will be available openly and shared under an Open Data Commons Open Database License ODbL (https://opendatacommons.org), which allows for commercial use. In order to widen re-use of database contents and creative products (i.e. innovations, publications) PRECAST project partners might use appropriate Creative Commons (CC) copyright licenses (https://creativecommons.org). This will allow others to copy, distribute and use data in their work while ensuring data owners get the credit they deserve for their work.

For scientific publications that are published in open access journals, the licensing schema of the publication repository should be followed. In most cases this is CC-BY schema thus PRECAST partners will strive to adopt it also for collected/generated data when depositing in public repositories. Creative Commons licensing schema CC-BY is the most accommodating of licenses offered and recommended for maximum dissemination and use of licensed materials because it lets others distribute, remix, tweak, and build upon licensed work, even commercially, as long as authors are credited for the original creation.

3.4.2. Data availability

Publishing in peer-reviewed scientific journals will be used as the main means of public disclosure of collected/generated data. All researchers will ensure that the data underlying peer-reviewed publications will be retained and made available for verification purposes after publication which is in accordance with accepted codes and standards. Whenever suitable, researchers will aim to publish in Open Data Journals (the list of examples could be found following the link: https://www.fosteropenscience.eu/taxonomy/term/114) and after publication underlying data will be uploaded into suitable public repository and stored in institutional servers.

All data ought to be disseminated through peer-reviewed publication will be made available at the time of publication or the latest after embargo period prescribed (up to 6months after publication) by the journal if the green open access of the paper is implemented. For data decided to be commercialized an embargo period could be foreseen until the patent application is resolved (typically 12months).

PRECAST consortium does not plan to restrict data availability and re-use after they are made available through peer-reviewed publications or public repositories. PRECAST consortium aims to keep collected data as long as possible on institutional storage. This will provide a database suitable for analysis of long-term trends which are important products that PRECAST partners will communicate to end users. All

generated data will remain stored for at least 10 years after they were collected/generated in order to allow results verification.

3.4.3. Use of data by third parties

PRECAST partners will predominantly re-use data coming from open sources and thus their use by third parties during implementation of PRECAST project and after depends solely on the decisions of the owners (See Table 2 for indicative list).

The majority of data will be available openly and shared under CC-BY license, which allows for commercial use by third parties. In case data providers, such as SMEs, industry or other organizations, do not wish to share their data a revenue model will be imposed.

3.3.4. Data quality assurance

Whenever possible, approved and verified methodology will be used for collecting/generating data. All data is collected by following the best practice in the respective field. The authors of the collected/generated data are responsible to ensure quality of the data before storage to institutional servers.

4. Allocation of resources

4.1. Costs form making data FAIR

4.1.1. Peer-reviewed publications (including uderlying data)

There are limited funds for open access charges available in the PRECAST budget. Therefore, consortium partners will be 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 green open access. 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 embargo period is too long, PRECAST consortium might opt for the gold open access.

To ensure Open Access after the lifetime of the PRECAST project, PRECAST consortium will, whenever possible, use free tools that are compatible with the requirements by Pilot on Open Research Data in Horizon 2020 (e.g. deposition of data and publications in OpenAIRE's Zenodo centralized repository that allows for free and long term deposition.

4.1.2. Responsibilities for data management

Responsibility for implementing principles defined by the Data Management Plan:

- Adding metadata - researchers that collect/generate and re-use data;
- Ethical issues - appointed Ethical Advisor;

- Deposition in appropriate repositories - researchers that collect/generate and re-use data;
- Enabling Open Access - researchers that collect/generate and re-use data.

The compliance to defined data management principles is monitored by Quality Manager and reported to Project Coordinator.

5. Data security

5.1. Secure storage and data recovery

PRECAST beneficiaries guarantee that all data collected during the project will be kept secure and unreachable by unauthorized persons. Data will be handled with appropriate confidentiality and technical security. Furthermore, each beneficiary will be responsible for the security of the research data at their institutions through the lifetime of the project.

PRECAST consortium partners will store all generated/collected data on a dedicated Data Storage System with dual controllers and dual power supply, or equivalent. Everything stored on those machines are copied on at least three Hard Disk Drives (HDD). In case of failure of one of the HDDs, data is secured on two others and within 24 hours the replacement HDD is obtained from the manufacturer. In case of electricity cut offs, dual power supply enables continuum by automatically swapping from electric network to UPS with diesel aggregate.

As an example, the data stored in the BioSense Institute Data Storage System are not exposed directly to the end users/internet thanks to two-line defense 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.

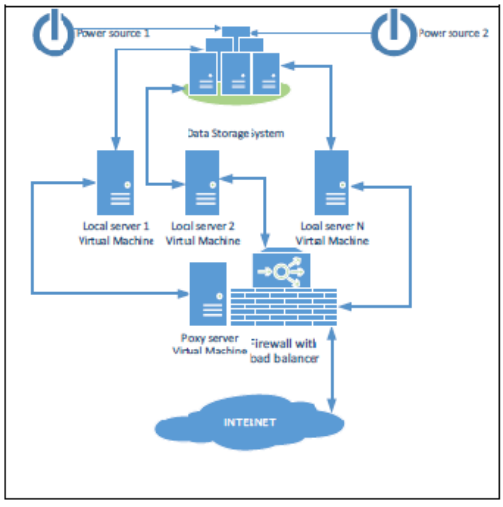

Figure 1: Architecture of the BioSense Institute Data Storage System

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.

5.2. 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 downloaded directly as web-based services from servers. In any case 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, WhatsUp).

Prior to 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.

In specific cases, when information related to members of project teams are shared, related to their expertise and research with a purpose to attend to trainings and workshops organized within the project, it is allowed to send personal information between partner institutions

All documents marked as "Confidential" will be available only to the partner who receives the information, and their further sharing and management are regulated with the Project Consortium agreement.

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 Horizon 2020 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).

5.3. Curation, preservation, and dissemination

5.3.1. Scientific publication

Due to the multidisciplinary nature of the PRECAST project, a connection to OpenAIRE's **Zenodo** (https://zenodo.org/) centralized repository will be established for the purpose of long-term duration,

preservation and dissemination of the scientific publications and underlying data. Zenodo specifically targets data and publications from EU projects and ensures discoverability, accessibility, and intelligibility over a wide range of subjects/themes. It supports Closed, Open and Embargoed[4] Access. However, only Open Access uploads are displayed on the Zenodo website front-page. Closed Access uploads are still discoverable through search queries, their DOI and any community collections where they are included. Metadata is licensed under Creative Commons "No Rights Reserved" license (CC-0)[5], except for email addresses. All metadata is exported via Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)[6] and can be harvested. Access to metadata and data files is provided over standard protocols such as HTTP and OAI-PMH. Zenodo accepts data under a variety of licenses, but extra benefits, in terms of visibility and credit, and additional services and upload quotas are offered to data deposited under the most open licenses.

Note: The full set of Zenodo "Terms of Use" can be found online at https://zenodo.org/terms while the full set of Zenodo Policies can be found online at https://zenodo.org/policies.

In case of closure of the repository, best efforts will be made to integrate all content into suitable alternative institutional and/or subject based repositories.

6. Ethical aspects

According to the Article 34 of the Grant Agreement, the beneficiaries must carry out the action in compliance with ethical principles (including the highest standards of research integrity) and applicable international, EU and national law.

In particular and relevant for the DMP, taking into account fundamental principles of reliability,

honesty, respect and accountability, the beneficiaries must ensure that persons carrying out research tasks:

- use techniques and methodologies (including for data collection and management) that are appropriate for the field(s) concerned.

[4] Users may deposit content under an embargo status and provide and end date for the embargo. The repository will restrict access to the data until the end of the embargo period; at which time, the content will become publically available automatically.

[5] Creative Commons "No Rights Reserved" license (CC0) enables scientists, educators, artists and other creators and owners of copyright- or database-protected content to waive those interests in their works and thereby place them as completely as possible in the public domain, so that others may freely build upon, enhance and reuse the works for any purposes without restriction under copyright or database law. Further information can be found at https://creativecommons.org/about/cc0.

[6] The Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) is a low-barrier mechanism for repository interoperability. Further information can be found at: https://www.openarchives.org/pmh/.

- allow — in addition to the open access obligations under Article 29.3 as much as possible and taking into account the legitimate interest of the beneficiaries — access to research data, in order to enable research to be reproduced.
- refrain from practicing any form of plagiarism, data falsification or fabrication.

As mentioned, consortium members are obliged to comply with international, EU and respective national laws. In addition, they must also adhere to their own institutional policies and procedures for data management, ethical standards, and codes of conduct, as well as national laws and guidelines for responsible research.

7. Legal aspects

Any collection and/or analysis of any personal data by the PRECAST consortium will be assessed by project's Ethical Advisor in order to ensure compliance with both the national Law on personal data protection (Official Gazette of RS no.97/2008, 104/2009, 68/2012 and 107/2012) and the EU General Data Protection Regulation (GDPR)[7].

8. PRECAST Data Management Templates

8.1. Data collection (questionnaire)

Describe what the purpose of the data collection/generation is and its relation to the objectives of the Task (i.e. what this data is required for)?

Describe what types of data the Task will generate/collect (e.g. bibliographic data, modeling data, survey replies, etc)

Will the task re-use any existing data and how?

[7] *http://ec.europa.eu/justice/data-protection/reform/files/regulation_oj_en.pdf*

What is the expected size of the data?

To whom it might be useful („data utility“)?

Where and how are the data stored?

What are the risk for the data?

8.2. FAIR Data (Findability, Accessibility, Interoperability, Reusability)

8.2.1. Making data FINDABLE

Are the data produced and/or used in the Task discoverable with metadata?

Does the task plan to make use of persistent and unique identifiers?

What naming convention do you follow?

8.2.2. Making data ACCESSIBLE

Which data produced and/or used in the Task will be made openly available as the default?

How will the data be made accessible (e.g. by deposition in a repository)?

What methods or software tools are needed to access the data?

Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.

If there are restrictions on use, how will access be provided?

8.2.3. Making data INTEROPERABLE

Is the Task allowing data exchange and re-use between researchers, institutions, organisations, countries, etc.?

What data and metadata vocabularies, standards or methodologies will the Task follow to make the data interoperable?

Will your task use standard vocabularies for the data types present in the data set to allow inter-disciplinary interoperability?

8.2.4. Making data RE-USABLE

How will the data be licensed to permit the widest re-use possible?

e.g. Creative Commons license CC-BY or CC-0 (according to the H2020 guidelines)

Be aware there are different licenses for research data (in comparison with publications), depending on the nature of these data (origin)

When will the data be made available for reuse?

How long is it intended that the data remains re-usable?

e.g. at least 10 years, for the lifetime of the repository

Are the data produced and/or used in the Task usable by the third parties, in particular after the end of the pilot/project? If the re-use of some data is restricted, explain why.

How is the data quality assured? Are data quality assurance processes described?

8.2.5. Data Security

What provisions are in the place for data security? (including data recovery as well as secure storage and transfer of sensitive data)?

Is the data safely stored in certified repositories for long term preservation and curation?

8.2.6. Allocation of resources

What are the costs of making data FAIR in your project? And how will these be covered?

e.g. Long-term storage, journal open access; project's budget

Who will be responsible for data management in the Task?

What are the costs of long-term preservation? And who decides how and what data will be kept and for how long?