TRUSTAI

TRANSPARENT, RELIABLE
& UNBIASED SMART TOOL

D8.2. DMP – Data Management Plan

March 31, 2021

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant agreement No 952060.

DOCUMENT CONTROL PAGE

DOCUMENT	**D8.2 – Requirements specification for the healthcare use case**
TYPE	**Report**
DISTRIBUTION LEVEL	Public
DUE DELIVERY DATE	31/03/2020
DATE OF DELIVERY	31/03/2020
VERSION	0.3
DELIVERABLE RESPONSIBLE	INESC TEC
AUTHOR (S)	Fábio Moreira (INESC TEC), Gonçalo Figueira (INESC TEC), Yulia Karimova (INESC TEC)
OFFICIAL REVIEWER/s	Gonçalo Figueira (INESC TEC)

DOCUMENT HISTORY

VERSION	AUTHORS	DATE	CONTENT AND CHANGES
0.1	Yulia Karimova (INESC TEC)	10/02/2021	First draft
0.2	Fábio Moreira (INESC TEC)	20/03/2020	Second draft considering partners contributions. Version submitted to reviewers.
0.3	Gonçalo Figueira (INESC TEC)	31/03/2020	Final revision

DISCLAIMER:

The sole responsibility for the content lies with the authors. It does not necessarily reflect the opinion of the CNECT or the European Commission (EC). CNECT or the EC are not responsible for any use that may be made of the information contained therein.

Table of Contents

Abbreviations and Acronyms

AI	Artificial Intelligence
EC	European Commission
EU	European Union
HCXAI	Human-centered Explainable AI
KPI	Key Performance Indicators
MS	Milestones
PM	Person Month
PR	Press Release
SMEs	Small and Medium-sized Enterprises
WP	Work Package
XAI	Explainable Artificial Intelligence

1. General information

Creators of DMP	• Yulia Karimova (INESC TEC/ FEUP) • Gonçalo Reis Figueira (INESC TEC) • Pedro Amorim (FEUP / INESC TEC) • Fábio Neves Moreira (INESC TEC) • Cristiane Ferreira (INESC TEC)
Affiliation	INESC TEC
Funder	European Union's Horizon 2020 Research and Innovation program under the Grant Agreement (GA)
Template	Horizon 2020
ORCID	• Yulia Karimova (https://orcid.org/0000-0002-1015-6709) • Gonçalo Figueira (https://orcid.org/0000-0001-6696-824X) • Pedro Amorim (https://orcid.org/0000-0001-7857-3405) • Fábio Moreira (https://orcid.org/0000-0003-4296-8207) • Critiane Ferreira (https://orcid.org/0000-0002-8680-1938)
Grant Number	952060
Project duration	01.10.2020 – 30.09.2024 (4 Years)

1.1. Project abstract

Artificial intelligence is single-handedly changing decision-making at different levels and sectors in often unpredictable and uncontrolled ways. Due to their black-box nature, existing models are difficult to interpret, and hence trust. Explainable AI is an emergent field, but, to ensure no loss of predictive power, many of the proposed approaches just build local explanators on top of powerful black-box models. To change this paradigm and create an equally powerful, yet fully explainable model, we need to be able to learn its structure. However, searching for both structure and parameters is extremely challenging. Moreover, there is the risk that the necessary variables and operators are not provided to the algorithm, which leads to more complex and less general models. It is clear that state-of-the-art, yet practical, real-world solutions cannot not come only from the computer science world. Our approach therefore consists in involving human intelligence in the discovery process, resulting in AI and humans working in concert to find better solutions (i.e. models that are effective, comprehensible and generalisable). This is made possible by employing 'explainable-by-design' symbolic models and learning algorithms, and by adopting a human-

centric, ‘guided empirical’ learning process that integrates cognition, machine learning and human-machine interaction, ultimately resulting in a Transparent, Reliable and Unbiased Smart Tool. This proposal aims to design TRUST, ensure its adequacy to tackle predictive and prescriptive problems, and create an innovation ecosystem around it, whereby academia and companies can further exploit it, independently or in collaboration. The proposed ‘human-guided symbolic learning’ should be the next ‘go-to paradigm’ for a wide range of sectors, where human agency / accountability is essential. These include healthcare, retail, energy, banking, insurance and public administration (of which the first three are explored in this project).

2. Data summary

In this project, we aim at learning directly a human-comprehensible model. We focus on machine learning applications which use images, text data, tabular data, including classification/regression. We want to improve the performance-explainability trade-off, so ML would be to combine the performance of (deep) neural networks and the explainability of theory-based approaches. We will mainly use approaches considering symbolic learning methods that work with analytical expressions that humans can understand and improve on. These methods will support a human-centric system, using theory and insight from humans to guide the machine’s empirical search. Humans are a key element of this project, as they need to interpret the output of the system and also guide the empirical search of machine learning. The resulting tool is TRUST – a Transparent, Reliable and Unbiased Smart Tool, that can be used in practice in a vast amount of applications and disrupt multiple sectors, where human control is essential. To achieve our vision, we propose a disruptive, but viable, paradigm with the objective to reach a breakthrough, rather than incremental improvements. To that end, we mobilise a multidisciplinary consortium (detailed information about consortium on section 3 of this DMP), joining experts from computer science, neuro and behavioural science, and industrial engineering, which have recognised international leadership in their own areas.

The objectives of this project are balanced between:

I. the development of the novel paradigm,
II. its exploration in different application domains, and
III. the creation of an innovation ecosystem around the TRUST human-centric framework.

Moreover, we integrate three key areas:

1. Machine learning, to search and provide explainable expressions, with as few task-specific hard-coded constraints as possible, and to allow the user to easily tweak the explainability;
2. Cognition, to incorporate behavioural and cognitive models of human explainability and causal reasoning in providing contrastive explanations (why P rather than Q?) and connecting causes;
3. Human-machine interaction, to establish trust and promote cooperation, and to interactively adjust the explainability of the ML model and its solutions, depending on the user’s level of expertise.

The project has three real-world use cases with different types of data, methodology, and management rules. Each use case will start by providing a toy problem, with core features, and then be progressively extended, as human experts validate the models and explanations generated. So, this DMP includes all necessary related Research Data Management (RDM) issues about each case separately. Moreover, this project includes 5 EU countries: Portugal, Estonia, France, Netherlands, Cyprus and 1 non-EU country: Turkey. Thus, this DMP also includes detailed information about data management related to the Turkish partner (see this on section 3 of this DMP).

2.1. Use cases

2.1.1. Cancer treatment (Healthcare)

This use case includes collaboration between Estonia, France, The Netherlands, Portugal, and Turkey, where the Dutch entity, CWI, is responsible. The use case will be developed at Leiden University Medical Center (LUMC). At the time of diagnosis, a choice between no intervention and treatment needs to be made. Treatment of paraganglioma can be done either by surgery or radiotherapy. However, it is not always necessary to use these treatments on the (most often benign) tumors. This is the case if the tumors stop growing and do not cause symptoms. Whether treatment is needed somewhere in the future and when that will be, is most often uncertain up front. Therefore, treatment is postponed until at least persistent growth of the tumor is demonstrated or until the tumor starts to cause (irreversible) symptoms. Forecasting the future development of paraganglioma allows us to make better decisions about the moment of treatment and follow-up. In order to support the physician's decision at each meeting with the patient, we plan to set the prediction tasks to be quantifiable and observable outcomes. Input to the prediction is all the information collected up to that time about the patient. The main challenge here is to forecast whether the patient will need treatment and if so, when. Key to this is likely being able to predict the growth of the paraganglioma. In addition, forecasting the probability of possible negative effects of the tumor such as hearing loss can help in better pinpointing the moment to proceed to treatment. The main users of the models will be the clinicians. However, since there is shared decision making between the clinicians and patients, the clinician needs to be able to explain these models to the patient. The use case aims at delivering the desired predictions through explainable AI models that clinicians can accept, validate, and use in clinical practice. To develop TRUST approaches, several data types will be used, including images, tumor-related data, signs and symptoms, DNA mutations, biochemical screenings, past treatments, and other generic features (e.g., sex, medication, smoking, alcohol, etc). The raw data collected by the hospital during the time of the project can be related to users of the health system (doctors, patients, technicians) and needs to be kept at the hospital, it should not be shared. Only samples of anonymized and masked data can be shared, with the sufficient conditions for the TRUST approaches (algorithms and interfaces) to be developed. In sum, it is to be ensured that no connection between the data rows and real individuals or entities should be possible. Publications using the processed data (anonymised, codified or pseudonymized) will be decided during the project, probably in the middle of the project. At the moment, it is not possible to define further details about the adopted software, dataset sizes, and publications reach but this information will be updated later in the project.

This use case will also be using the informed consent procedures. Templates of the informed consent/ assent forms and information sheets (in language and terms intelligible to the participants) will be created later and will be approved by the Ethics Committee. If use case 1 will involve also clinical aspects, the following documents/information must be created: (i) Final version of study protocol as submitted to regulators/ethics committee(s), (ii) Registration number of clinical study in a WHO-or ICMJE- approved registry (with the possibility to post results), (iii) Approvals (ethics committees and national competent authority if applicable) required for invitation/enrolment of the first subject in at least one clinical centre. Copies of pertinent opinions/approvals by ethics committees and/or competent authorities for the research with humans also will be created and kept on file. The outcome of future medical ethical committee submissions will be shared, and, as is common practice at LUMC, be kept on file. More detail about informed consent can be seen at Grant Agreement. The responsibles for data management of these data are Peter Bosman at CWI and Tanja Alderliesten at LUMC and at CWI.

2.1.2. Time slot selection (Online retail)

This use case includes collaboration between four countries, Estonia, France, Portugal, and Turkey. Three Portuguese entities, INESC TEC, Sonae MC, and LTPlabs are involved, where the latter is responsible. This use case is focused on online retail of different industries, including fashion, electronics and grocery. More concretely, it is focused on decisions to be made on the fly (i.e., in real-time), such as selecting delivery time slots to be offered, from where items will ship and by what shipping method. Logically, AI systems are necessary to make these decisions in an optimized way and these decisions need to be based on and simulated using historical data, in order to achieve and assess the fitness of each expression. During this use case the data that will be used is generally not sensitive and not personal. The most sensitive information could be the composition of the shopping basket and the address location. However, customers' privacy will be ensured by their anonymisation. The first type of data is related to customers (demographic data, shopping history, current basket, destination, customer profile, purchasing habits) and it will help to predict how much each customer is willing to pay to have his/her order arriving within a specified delivery window and to extract insight on customer preferences. The second type of data concerns logistics (logistics load, planned routing, delivery conditions, logistics costs) that will help to accurately estimate the cost of serving a customer within a given time window and to the ability to provide a cost breakdown. In this use case, we will involve a large grocery retailer, Sonae MC, which has more than 200 brick-and-mortar stores, and a considerable volume of online orders (and which has committed to providing all the necessary raw data). We focus on the time slot selection problem since it has a huge impact on the optimization of last-mile deliveries. In addition, when selecting time slots, there is a trade-off to be considered between customer satisfaction (i.e., offering the most appropriate slots for a given customer profile) and operational efficiency (i.e., aggregate orders in certain slots so as to optimize transportation). This process also involves multiple stakeholders in the company, namely marketing and operations divisions, though no data involving these stakeholders is necessary. Raw data concerning retail customers are sensitive, personal and private so they cannot be open publicly, but the publishing of the processed data (anonymised, codified or pseudonymized) will be decided during the project, probably in the middle of the project. Other partners should not need the raw data related to this use case. In case any partner needs realistic data to develop algorithmic approaches or realistic cases to test the prospective interfaces, the data will be anonymised and masked so that no connection exists

to a real person or entity. The responsible for data management of these data is Francisco Amorim (LTPlabs). The ownership of these data is pertence to Sonae MC. The raw data will never be shared, only the realistic samples generated from anonymised data can be shared or published (i.e. as part of a scientific publication). The type of data that can possibly be published will be defined upon the date of the refereed publication.

2.1.3. Demand forecast (Energy)

This use case includes collaboration between Cyprus, France, The Netherlands, Portugal, and Turkey, where the Crypriot entity, Apintech, is responsible. This use case is focused on reading and processing, in real-time, the data from all required data sources, related to the energy in the building, in a trustful way to train a forecasting ML model and use it subsequently in daily, hourly and minute demand prediction. In this TRUST use case, we will focus on a single building (that can provide essential real-time energy/ indoor quality data). Data about non-linear factors (weather, indoor conditions of a building, behavioural data) will be collected to help on the performance and accuracy of electrical energy consumption forecasting, using a hybrid between AI and SI -swarm intelligence- methods. Moreover, it will be collected data about past customer data (use/device), profiles (consumptions), type of building (residential, non-residential); time interval (per minutes, hour, month, year); electrical load (historical load); weather data (temperature, dry bulb temperature, dew point temperature, wet point temperature, air temperature, humidity, wind speed, wind direction, brightness of sun, precipitation, vapor pressure, global/sola radiation, sky condition); indoor data (ambient temperature, electric usage, occupancy); and other data (thermal index, calendar, size of house, the living standard, social development, urban construction and development, the natural condition, heat transfer, size of building). These data will be used on the development of a suitable and open data API (in progress). Note that no personal or sensible data is involved in the open API however, this use case is likely to involve further processing of previously collected personal data (secondary use) from buildings residing in the EU. Again, the consumer's privacy will be ensured, by anonymisation/pseudonymisation, using the best practices. Since the necessity to transfer data to other partners is present, anonymised samples will be available to be shared. The idea is to provide TRUST partners with realistic data so that xAI algorithms and interfaces can be developed with sufficient degree of realism. All the datasets shared will ensure that no connection between real persons or entities can be established. Good quality datasets will be critical to the main objective of this use case which is to develop an advanced and explainable energy forecasting tool for city building and also for entire cities. Furthermore, each energy consumption forecast is supposed to be disaggregated into a more detailed form, increasing the granularity of the forecasts and guiding the users by providing causal rules and counterfactuals that can be used to support decision making. Multiple interfaces are to be developed for each type of user (e.g., building owners and policy makers). At the moment, it is not possible to define further details about the adopted software, dataset sizes, and publications reach but this information will be updated later in the project. The owner of the raw data considered in this use case is Apintech.

2.2. TRUST outputs

All the outputs of the project are to be preserved under internal and trustable storage platforms (emails, DPIA, Consent informs, agreements, Communication & Dissemination

Plan (CDP), NDA non-disclosure agreements, research code, research papers). Part of the data and outputs of the project can be useful to reuse in different sectors such as healthcare, retail and energy, banking, insurance and public administration, since decision making in these contexts must be transparent and explainable. In other words, the data is very important to markets with great societal and economic potential. Moreover, the development, implementation and exploitation of TRUST in those sectors must be supported in a thorough discussion of the organizational, societal, ethical and legal implications that these tools might have. This should also contribute to the wider debate of AI systems, decision support (e.g. medical prescription, operations management, financial investments), decisions that affect people (e.g. whether customers or applicants are being treated fairly), and accountability providence to external entities (e.g. regulatory bodies, citizens).

3. *FAIR* data

3.1. Use cases

3.1.1. Cancer treatment (Healthcare):

The raw data of this use case will not be publicly open, not accessible to any partners, neither for request. The raw data belongs to LUMC, and it will be preserved according to their needs and rules which are out of the scope of TRUST. Although samples of processed data will be available to all partners, only some of this data will be published. The conditions in which the publishable datasets will be revealed in the next DMP.

3.1.2. Time slot selection (Online Retail):

The raw data of this use case will not be publicly open, not accessible to any partners, neither for request. The raw data belongs to Sonae MC, and it will be preserved according to their needs and rules (with LTPlabs) which are out of the scope of TRUST. The conditions in which the publishable datasets will be revealed in the next DMP.

3.1.3. Demand forecast (Energy):

The raw data of this use case will not be publicly open, not accessible to any partners, neither for request. The raw data belongs to Apintech, and it will be preserved according to their needs and rules which are out of the scope of TRUST. However, a real-time source of data is planned to be publicly available. The public data figuring in this API will be processed data, ensuring that no personal or sensible data is publicly available. Additional processed data is to be shared with all partners of TRUST. The conditions in which the publishable datasets will be revealed in the next DMP.

3.2. Repositories

After the definition of the data to be published, the data repository that will be used will be defined. The Zenodo repository (http://www.zenodo.org/) is one of the suggested repositories to data publish, and to open other type of data specified in the DMP. Moreover, for each dataset deposited on Zenodo, a DOI will be defined to make the data more FAIR and citable. Moreover, Zenodo allows version control of datasets. The published datasets are to be kept in Zenodo for an indefinite period of time in order to support other research projects and comparisons. INESC TEC research data repository (https://rdm.inesctec.pt/) is the second repository suggested to data publishing. It also allows the definition of a DOI for each dataset and provides more flexibility on the description of the data.

4. Allocation of resources

4.1. Costs for making the data *FAIR*

To make the data FAIR it is necessary to have access to the chosen research data repositories. Zenodo and INESC TEC repositories are free and open-access, thus no additional costs are required. The processed data that will be stored at INESC TEC cloud drive is also freely accessible to all the partners of the project. The scientific publications resulting from TRUST will also be shared at INESC TEC drive. Recall that, these publications are to be preferably published in open-access journals but if that is not the case, public versions should be provided through INESC TEC drive links.

In sum, during the project the following assets will be used:

- Hardware/devices: work desktop, laptop computers, and institutional servers (costs of this equipment is detailed in the Grant Agreement document).
- Software: Windows, Linux, Office 365, Visual Studio.
- Cloud services: https://drive.inesctec.pt, with access granted to all partners through provided links of shared folders. Maintenance activities related to INESC TEC drive are to be carried out by the INESC IT structure.

Any updates on the aforementioned information will be added in the next versions of this DMP.

4.2. Responsibilities for data management in the project

A multidisciplinary Consortium was created with the participation of experts from R&D organisations (INESC TEC, INRIA, NOW-I, CITIS), universities (University of Tartu, LUMC) and ambitious SMEs with different profiles: Apintech works with real-time, big data collected from sensors and IoT, LTP conducts analytics-based consulting, and Tazi AI commercializes an explainable AI platform. The organisational structure of the Consortium includes an independent Ethics and Data Protection Board (EDPB), composed of elements of the different institutions, and necessarily including those involved in the use cases (INESC

TEC/LTPlabs, CWI/LUMC and POLIS21). EDPB includes a Data Protection Officer from each WP leader and pilot:

- UC1: LUMC data protection officer, infoavg@lumc.nl
- UC2: Vasco Dias, dpo@inesctec.pt
- UC3: Costas Daskalakis, info@apintech.com

This consortium body will monitor the ethics issues involved in this project and how they are handled.

Additional responsibilities include:

- Responsible for DMP creation: Gonçalo Reis Figueira (https://orcid.org/0000-0001-6696-824X), Fábio Neves Moreira (https://orcid.org/0000-0003-4296-8207), Yulia Karimova (http://orcid.org/0000-0002-1015-6709)
- Responsible for DPIA, Grant Agreement, and other legal documents. creation: Vasco Dias (INESC TEC)

Responsible for the RDM in each use case:

1. **Cancer treatment (Healthcare):**

- Responsible for the collection of the data: Peter Bosman (CWI)
- Responsible for the processing and preservation of the data: Peter Bosman (CWI)
- Responsible for backups: Peter Bosman (CWI)
- Responsible for publishing and sharing data: Peter Bosman (CWI)

2. **Time slot selection (Online Retail):**

- Responsible for the collection of the data: Francisco Amorim (LTPlabs)
- Responsible for the processing and preservation of the data: Francisco Amorim (LTPlabs)
- Responsible for backups: Francisco Amorim (LTPlabs)
- Responsible for publishing and sharing data: Francisco Amorim (LTPlabs)

3. **Demand forecast (Energy):**

- Responsible for the collection of the data: Nikos Sakkas (Apintech)
- Responsible for the processing and preservation of the data: Nikos Sakkas (Apintech)
- Responsible for backups: Nikos Sakkas (Apintech)
- Responsible for publishing and sharing data: Nikos Sakkas (Apintech)

4.3. Potential value of long-term preservation

The data acquired, collected, and generated during this project are unique and thus very important for other researchers as well as for educational purposes. Some data will be preserved (will be defined later) to enable initially unforeseen uses of the data and to guarantee fully-documented and reproducible data from the project, ensuring the reuse of the data in multiple domains and sectors such as healthcare, retail and energy, banking, insurance and public administration, and different applications.

Some data will be preserved at the Zenodo repository without any limitation after the project without any embargo period (more detail will be described in the following DMP versions).

The INESC TEC IT structure will be responsible for any action related to the long-time preservation in accordance with the repository guarantees, that also will be chosen later. The corresponding information will be added if necessary in a further version of the DMP.

Each preserved dataset of the project will have Digital Object Identifiers (DOIs) attributed by the repository (Zenodo or INESC TEC).

5. Data security, access, storage and backups

Different criteria have been defined according to each use case:

1. **Cancer treatment (Healthcare):**

Before the beginning of work in use case 1, all pertinent ethics committee opinion/authorisation will be submitted as a report by the EDPB. The raw data never will pass to other partners from LUMC, so the preservation rules, access to raw data will be defined by LUMC and described with more detail on the further DMP. The measures to protect them will be through following a good clinical practice. Moreover, LUMC will anonymize the data before sharing it with CWI, where the data will be processed. Minimizing the risk of identifying any individuals as much as possible.

In this use case, the raw data will be preserved on the database of the LUMC hospital and it will not be shared with others. Only processed data will be shared with others through the https://drive.inesctec.pt. However, the data that will be shared is not defined yet. The backups of the processed data will be managed by the INESC TEC drive maintainers. These backups are usually preserved for more than one year. All the data that is available at the INESC TEC drive will also be accessible to all partners of the TRUST AI project. To access and process data, licensed software will be used to reduce the risk of intruders. Moreover, all the computers have properly updated antivirus software. The browsers have Adblock installed that permit block the pop-up pages and other insecure connections. All technical issues related to the software will be controlled by each member of the project and in case of necessary support, they contact the IT staff of the responsible entity. The paper documents are not to be preserved unless they are indispensable to continue the project or to serve as a proof of any kind.

2. **Time slot selection (Online Retail):**

In this use case, the raw data provided by Sonae MC will be preserved on a database owned by LTPlabs and will not be shared with others. Only processed data will be shared with others through the INESC TEC cloud service available at https://drive.inesctec.pt. The data that will be shared will depend on the necessities of each of the involved partners, yet those data will always be processed (anonymised and masked) so that no connection between real persons or entities can be defined. The raw data will have periodic backups according to the internal

policies of LTPlabs. The backups of the processed data will be managed by the INESC TEC drive maintainers. These backups are usually preserved for more than one year. All the data that is available at the INESC TEC drive will also be accessible to all partners of the TRUST AI project.

To access and process data, licensed software will be used to reduce the risk of intruders. Moreover, all the computers have properly updated antivirus software. The browsers have Adblock installed that permit block the pop-up pages and other insecure connections. All technical issues related to the software will be controlled by each member of the project and in case of necessary support, they contact the IT staff of the responsible entity. The paper documents are not to be preserved unless they are indispensable to continue the project or to serve as a proof of any kind.

3. **Demand forecast (Energy):**

In this use case, the raw data will be preserved on Apintech databases and it will not be shared with others. Only processed data will be shared with other partners through an open API (in real-time) and through the https://drive.inesctec.pt. The processed data that will be shared is not defined yet. The raw data will not have backups accessible to the partners of TRUST. The backups of the processed data will be managed by the INESC TEC drive maintainers. These backups are usually preserved for more than one year. All the data that is available at the INESC TEC drive will also be accessible to all partners of the TRUST AI project. To access and process data, licensed software will be used to reduce the risk of intruders. Moreover, all the computers have properly updated antivirus software. The browsers have Adblock installed that permit block the pop-up pages and other insecure connections. All technical issues related to the software will be controlled by each member of the project and in case of necessary support, they contact the IT staff of the responsible entity. The paper documents are not to be preserved unless they are indispensable to continue the project or to serve as a proof of any kind.

The raw data will not be transferred from a non-EU country to the EU, neither in the case related to Turkey. If any processed data transfer occurs, all appropriate documentation will be provided and described in the new version of the DMP.

All technical and organisational measures will be described on DPIA according to art. 35º GDPR. In further versions of the DMP this information will also be added. Moreover, it will be described the security measures that will be implemented to prevent unauthorised access to personal data or the equipment used for processing, anonymization/pseudonymisation techniques and data protection policies.

More detail about right access can be read in the Consortium and Grant agreements.

6. Ethical aspects

The TRUST-AI Project is financed by Horizon 2020 (Nº 952060), and responds to all existing requirements related to the Research Data Management and Protection of Personal Data. All partners follow the Regulations defined by the EU, and other agreements defined by the Consortium of this project which define a set of the rules to guide projects. This DMP will be verified by the INESC TEC DPO, will follow the General Data Protection Regulation and will

be monitored every year to add all changes that occurred in the project. In addition, DPIA will also be created. Moreover, was created a Grant Agreement and Consortium Agreement where described in detail all Ethics rules between the partners.