Foundations of Trustworthy AI – Integrating Reasoning, Learning and Optimization
TAILOR
Grant Agreement Number 952215

Data Management Plan, Report

Document type (nature)	Report
Deliverable No	D1.6
Work package number(s)	WP1
Date	Due February 28, 2021
Responsible Beneficiary	LiU, ID 1
Author(s)	Trine Vikinge
Publicity level	Public
Short description (Please insert the text in the Description of Deliverables in the Appendix 1 of the Amendment.)	"A Data Management Plan will be prepared to describe which data will be generated, collected and processed within the TAILOR project."

History & revision plan (month 18 & 36)			
Revision	**Date**	**Modification**	**Author**
1.0	2021 February 28	First version	Trine Vikinge

Document Review		
Reviewer	**Partner ID / Acronym**	**Date of report approval**
Marc Schoenauer	#3, Inria	2021-02-20
Maurizio Lenzerini	#6, UNIROMA1	2021-02-23

For review, the template provided in Folder B on Drive has been used. The review documents are saved in the dedicated EMDESK folder.

Table of Contents

Summary of the report

The project is committed to be in compliance with applicable rules and regulations and with the highest moral standards in the area. The purpose of the project is building the capacity of providing the scientific foundations for Trustworthy AI in Europe by developing a network of research excellence centres leveraging and combining learning, optimization and reasoning. Lawfulness, ethics, and robustness are pillars of trustworthy AI. As a consequence, setting and living up to high standards in the area of data management is a cornerstone of the project.

Privacy
All TAILOR partners are committed to comply with applicable rules and regulations. This DMP is particularly concerned with compliance with the GDPR and with the FAIR principles. The GDPR prescribes that management of personal data is the responsibility of each legal actor in possession of such, and that each such actor is required to have a data protection officer (DPO).
Each partner organisation of TAILOR is thus obliged to have a data protection/data management policy for management of personal data (Local DMP), including routines for reviews and updates, as well as a DPO. Each partner will, when required according to the GDPR, do DPIAs (Data protection Impact Assessments), under due supervision of the organisation DPO. Ethical aspects of management of personal and personal sensitive data in TAILOR are described in the deliverables of WP13 Ethics requirements.

Should absence of required local policy documents be identified, the partner will be encouraged to rectify. Should a conflict result, the consortium conflict resolution routine as described in the agreements will be used to solve it.

Documentation assuring Local DMPs will be kept on file by the TAILOR project manager, and routines are set up to assure that compliance can be reasonably monitored.

FAIR
A part of the Open Research Data (ORD) pilot is concerned with making data management FAIR. In general terms, 'FAIR' means that data-sets are *findable*, *accessible*, *interoperable* and *re-usable*. The principles imply no restrictions with regard to specific technology, standard, or implementation-solution. TAILOR is a network with 54 partners including both universities and companies, and the technologies, standards, and solutions for assuring compliance with the FAIR principles will span a very broad range.

All TAILOR partners that are legal owners of data must assure that local data management plans or policies (Local DMP) that covers the FAIR principles are in place.

The TAILOR project manager will be responsible for assuring that all partners are aware and informed. Forms will be used to collect statements concerned with FAIR compliance and to identify the lack of such. The project manager is responsible for the process.

Partners are obliged to have local DMPs

All partners are expected to comply with all adequate aspects of the GDPR and the FAIR principles, including to have local DMPs in place that cover management of data in the TAILOR project. The Local DMPs shall cover management of both sensitive and non-sensitive personal data, and routines for data protection impact assessment (DPIA). When applicable, partners are expected to perform DPIAs in collaboration with the DPO of their own organisation. (Local DMPs may have other names.) Routines for compliance with the FAIR principles can be covered in the Local DMP or a different policy document of the organisation.

The checklist at the end of this document, Appendix 1, should be used to check that the Local DMP, or other policy document, covers the items included in the DMP template provided by the EC. (Link to the template and guidance are found in the list of references)

Routines implemented to simplify compliance

It is expected that all partners of TAILOR comply with this DMP, that is, with the Open Access policy, the GDPR, and with the FAIR principles.

1. All who are involved in the project will have access to and be encouraged to read this plan.
2. One representative of each partner, normally the one representing the partner in the General Assembly, will be asked to answer questionnaires based on the table in Appendix 2.
3. All those engaged in the project, and particularly the partner representatives of the General Assembly, are encouraged to engage in dialogues about data management in their task working groups. If non-compliance is identified or suspected, the Project manager shall be contacted without delay.
4. DMP compliance matters will be on the agenda of the WP-leader risk assessment meetings held twice per year.
5. An internal discussion forum and a repository of examples may be established.

All involved in TAILOR are expected to read and acknowledge the principles and guidelines explained by the EC in the document Ethics and Data Protection.

Introduction

TAILOR takes part in the pilot for open access to research data (ORD pilot) launched by the EC. The pilot aims to improve and maximise access to as well as-use of research data generated by Horizon 2020 projects, taking into account the need to balance openness and protection of scientific information, commercialisation and IPR, privacy concerns, security, data management and preservation questions, by means of implementation of the FAIR principles.

Part 1 of this plan is about our approach to Open data with regard to **publication** of articles. Part 2 of this plan is about our approach to Open data with regard to research **datasets** and the FAIR principles (findable, accessible, interoperable, and re-usable data).

Our approach to, and management of, ethical aspects of data, such as privacy and profiling, are described in greater detail in the deliverables of WP13.

Our approach and strategies for dissemination, protection, and exploitation is further elaborated in the dissemination plan deliverables of WP12.

Data management in network projects

The TAILOR project has 54 partners. All are engaged in contributing to trustworthy AI, by finding solutions to principle challenges concerned with combining learning, reasoning, and optimisation. The specific research areas of the project partners are diverse, and the questions to be answered in a DMP in line with the suggested template is neither feasible nor possible to be completed for the projects as a whole. Rather, data management must be governed locally, by each and all legal owners of any data applied in the project.

The categories of data expected to be managed in the project are described below in Part 2.

TAILOR is one of four networks of excellence (NoE) financed under H2020 ICT-48. We have consulted with the other NoEs as well as with the Coordination and Support Action (CSA) under the same call. Some of our partners are also partners of the other NoEs or the CSA, and by coordinating and aligning our strategies, we may simplify compliance and avoid introducing conflicts.

Open Access approach in TAILOR

Open access (OA) is about providing free of charge online access to scientific information for everybody; our communities and all EU citizens.

BOX 1

The Grant agreement signed by all TAILOR partners includes statements on Open Access (Article 29):

“Each beneficiary must ensure **open access (free of charge online access for any user) to all peer-reviewed scientific publications relating to its results.**

In particular, it must:

(a) as soon as possible and at the latest on publication, deposit a machine-readable electronic copy of the published version or final peer-reviewed manuscript accepted for publication in a repository for scientific publications;

Moreover, the beneficiary must aim to deposit at the same time the research data needed to validate the results presented in the deposited scientific publications.

(b) ensure open access to the deposited publication — via the repository — at the latest:

(i) on publication, if an electronic version is available for free via the publisher, or
(ii) within six months of publication in any other case.
(c) ensure open access — via the repository — to the bibliographic metadata that identify the deposited publication.
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.”

“Access” includes the right to read, download and print as well as the right to copy, distribute, search, link, crawl and mine. That is; to re-use.
“Scientific information” typically means

- peer-reviewed scientific research articles (published in scholarly journals), or
- research data (data underlying publications, curated data and/or raw data).

Figure 1 below schematically and simplified illustrates data flow, starting with input on the left hand side through the process of research to dissemination, publication, and deposition of data on the right hand side.

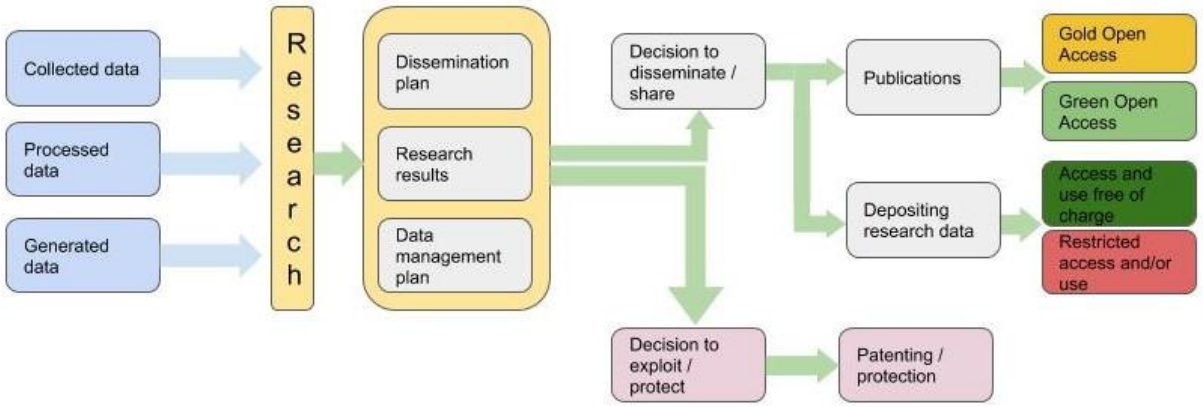

Figure 1: Schematic and simplified illustration of data flow, adapted from https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/open-access_en.htm

Part 1: Strategy for Publication

Aiming for free of charge access to publications

The two main routes to **open access to publications** are called “green” and “gold” open access. In green open access, the **author** deposits the published article / final peer-reviewed manuscript in an online repository. An example of such a repository is ArXive.org. In gold

open access, the article is immediately published in open access mode by the publisher. A fee, Article Processing Charge (APC), is paid by the authors to the publisher. Examples of such are scientific journals like Nature, Science, and those of the IEEE.

The open access requirements do not imply an obligation to publish results. But if publication is chosen as a means of dissemination, it becomes an obligation.

Not publishing may be desirable

Companies are heavily involved in TAILOR. Particularly in WP8, Industry, Innovation and Transfer program, AI challenges of industry are in focus. For actions involving industry research activities (e.g. via transfer labs, industrial internships, PhD-programmes) publishing outcomes may counteract goal achievement, and not making it openly accessible may be desired. All exceptions from open access needs to be motivated and should be duly documented in the WP deliverables.

Alternatives to publication are to keep the results confidential or to seek patent protection, as is illustrated in Figure 1. Any of these may be preferred if commercial exploitation is desirable. The limitations and restrictions stated by each partner in the Consortium Agreement, Attachment 1, with regard to Background are to be duly considered before dissemination. Background is defined as “data, know-how or information […] that is needed to implement the action or exploit the results”.

The dissemination strategy and considerations are further covered in deliverables of WP12 Dissemination and outreach.

All partners have signed off on the Open Access strategy as outlined in the GA. Respecting partners' need for restricting access shall always be weighted against the desire for openness.

BOX 2
GA, section from Article 29:
As an exception, the beneficiaries do not have to ensure open access to specific parts of their research data under Point (a) if the achievement of the action's main objective (as described in Annex 1) would be jeopardised by making those specific parts of the research data openly accessible.

WP leaders monitor compliance with Open Access policy

WP leaders will be responsible for monitoring open access strategy compliance in their own WPs. Should conflicts or disputes occur, WP leaders will present the case to the EB, and the conflict resolution procedure prescribed in the CA will be applied.

All exemptions from Open Access policy shall be duly motivated and reported in the appropriate WP deliverables.
Potential conflicts regarding open access to project outcome will be raised to the EB, and the conflict resolution procedures described in the CA are to be followed.

Part 2: Strategy for FAIR data management

Data management in TAILOR must strive towards being in compliance with the FAIR principles.

What are the FAIR principles?

The FAIR principles are:

- F Making data findable, including provisions for metadata
- A Making data openly accessible
- I Making data interoperable
- R Increase data re-use (through clarifying licences)

More about the principles can be found in this Nature article from 2016 The FAIR Guiding Principles for scientific data management and stewardship.

Box 3:
From GA, Article 29:
Regarding the digital research data **generated** in the action ('data'), the beneficiaries must:
(a) deposit in a research data repository and take measures to make it possible for third parties to access, mine, exploit, reproduce and disseminate — free of charge for any user — the following:

1. the data, including associated metadata, needed to validate the results presented in scientific publications, as soon as possible;
2. other data, including associated metadata, as specified and within the deadlines (to be specified in Local DMPs)

(b) provide information — via the repository — about tools and instruments at the disposal of the beneficiaries and necessary for validating the results (and — where possible — provide the tools and instruments themselves).

The deadlines for deposition of "other data", point 2 in the box above, will depend on local factors. It is recommended that deadlines are specified in the Local DMPs, and particularly that deviations from standard routines are duly motivated and documented.

Data categories in the project

Data managed in the project can be grouped as follows:

A. Personal (not sensitive) data collected in first hand for non-scientific purposes, such as
    a. in connection with network activities
    b. for connectivity fund applications,
    c. for summer schools
    d. from social media, i.e. visits on webpage, youtube channel etc
B. Data from repositories for scientific use
    a. from open access databases
    b. from restricted access databases
    c. provided by partner or by external actor
C. Data collected or generated for research purposes as part of the TAILOR project that may be
    a. sensitive personal data
    b. personal data
    c. other data

For all groups of data above, they should be managed lawfully and ethical, in compliance with EU regulations and local DMP and applicable policies.

Also, to the greatest degree possible, **data used for research**, independent of origin, shall be made as openly accessible as possible and deviations from open access shall be duly motivated and documented. Aspects of this pertaining to the different data groups are described below.

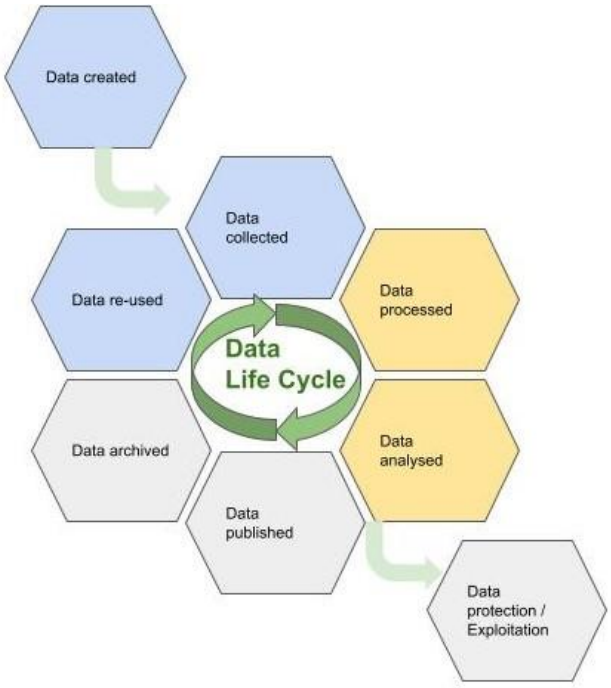

Figure 2: an illustration of data life cycle.

Figure 2 above is a simplified and schematic illustration of data life cycle and re-cycle which includes input of new data as well as output of data to be protected and/or exploited.

Making research data FAIR in TAILOR

The context and technical infrastructure, etc, of the project partners vary significantly, and for this reason one single FAIR policy is not feasible or meaningful. However, the partners agree that FAIR is a factor contributing to AI trustworthiness, i.e. a factor affecting positively the chances of reaching the project goals. All partners are expected to have policy documents for FAIR data management, either as separate documents/policies or as part of Local DMP. Below are some general statements on FAIR in TAILOR. For more details, partners' policies must be consulted.

The use of data items, even though made publicly available, is - as a rule - subject to certain license terms, such as Creative Commons (CC) licenses. The reproduction, making available and modification of copyrighted content qualifies as copyright relevant actions and, in principle, require authorization by the right holder, unless exceptions apply. Policies for compliance shall be included in local DMP or similar.

The basic rights and license terms will be attached and communicated via metadata. The licensing and access restrictions are - to reasonable extent - enforced by the technical infrastructure. Each partner is encouraged to include routines for this in their Local DPO. The TAILOR project partners are involved in the use of data from repositories, and likewise, the placing of their generated and processed data in such repositories. Partners are encouraged to comply with metadata conventions applicable for the discipline and to have routines for naming, keyword, routines for version management, etc.

For dataset identification, there is no one standard identification mechanism suitable to all data sets applied in the project. Persistent and unique identifiers such as Digital Object Identifiers (DOI) are used when possible, but partners may choose to use other means for unique identification.

Datasets resulting from research in the project shall in general be made openly accessible, as is described above, and exemptions shall be motivated and documented in the appropriate deliverable report.

In general, the project partners are pioneers in making data sets openly accessible, and will continue to be so in this project. All partners are responsible that the repositories they use for their data sets are suitable and appropriate for the purpose of contributing to the goals.

The (TAILOR) Trustworthy AI Repository will be considered as a permanent AI repository for relevant data sets. However, the partner owning the data is responsible to decide on what repository is appropriate to use. If specific methods or software tools are needed to access the data, required documentation and possibly open source code will be included with the dataset.

Generalisations of vocabularies, standards or methodologies with regard to interoperability are difficult to make. Partners will make efforts to facilitate interoperability, and this will be described in their local DMP or similar document.

Partners local DMPs or similar documents will describe local policies for how to increase reuse of datasets. In general, a policy is desirable that specifies how the data will be licenced to permit the widest reuse possible, and when the data will be made available for reuse. If applicable, if a period of data embargo is needed, this needs to be motivated and the embargo time specified. This is to be documented and filed in compliance with the local policies.

Also, partner policies should include whether the data produced and/or used in the project is usable by third parties, in particular after the end of the project. If the re-use of some data is restricted, this should be motivated and documented accordingly.

All partners are encouraged to have routines and processes for assurance of data quality, including routines covering aspects of the length of time for which the data will remain re-usable.

Management of personal data

Personal data are defined extremely broadly and include “any information relating to an identified or identifiable natural person”[1]. For personal data, both when generated in the project and when licensed, partners will give special attention to the confidentiality of data storage and processing. They will commit to implement all appropriate technical and organisational measures necessary in order to protect potential personal data against accidental or unlawful destruction or accidental loss, alteration, unauthorised disclosure or access, and against all other unlawful forms of processing, taking into account the particularity of the performed processing operations. Procedures for this are expected to be included in Local DMP.

Personal data in category A are personal data collected for non-scientific purposes, such as in connection with network activities, for connectivity fund applications, for summer schools, from social media, i.e. visits on webpage, youtube channel etc, and are regulated under the GDPR. Local DMP and policies include management of such data for administrative purposes.

If research is to be performed on such data, of category A, the procedures concerned with recruitment of external participants, described in the ethics deliverable D13.2, must be complied with.

[1] https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/ethics/h2020_hi_ethics-data-protection_en.pdf

In case previously collected personal data is to be further processed, the appropriate technical and organisational measures must be in place to safeguard the rights of the data subjects. Particularly, the Local DMP must include all required routines to assure regulatory and ethical compliance.

In general the research exception in Art. 5 (1) lit. (b) GDPR applies, which states that the further processing of personal data for research purposes is not considered incompatible with the original purpose for which the data had been collected as long as appropriate measures and safeguards for the rights of the data subjects are in place. The principle of data minimization and storage limitation also need to be seen with respect to the research goals. Art. 5 (1) (e) GDPR provides another exception for research.

With regard to category C data; Data collected or generated for research purposes, as part of the TAILOR project, that may be personal data, the ethics deliverables, D13, points to the need of informed consent and ethics authorities approvals, and particularly describes the routines implemented in the project to assure proper management of such. Particularly, if sensitive personal data is to be collected, the national ethics authorities' routines for obtaining approval of the research must be followed.

Even for personal data that is not of sensitive character, caution should be practiced. Article 35 of the GDPR states that:

> "Where a type of processing in particular using new technologies, and taking into account the nature, scope, context and purposes of the processing, is likely to result in a high risk to the rights and freedoms of natural persons, the controller shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data. A single assessment may address a set of similar processing operations that present similar high risks."

A Data Protection Impact Assessment, DPIA, shall be done if this paragraph applies, that is, if the processing to be carried out carries a risk of inflicting harm on individuals whose data is included. If in doubt, the organisation DPO shall always be consulted.

Researchers in TAILOR may develop tools where personal data (not sensitive data) is combined with publicly accessible data, primarily bibliographic data, for matching purposes.

In general, if personal data is anonymised or pseudonymised so it no longer relate to identifiable persons, such as aggregate and statistical data, or data that have otherwise been rendered anonymous, it is not considered to be personal data.

If tools developed in the project are to be made publicly available and if there is a risk that individuals can be identified, it is recommended that the local DPO is contacted and a DPIA considered.

When used in research, all personal data will be pseudonymised or anonymized by applying for example privacy-preserving methods (such as differential privacy, k-anonymity, randomization, etc.) that guarantees adequate privacy protection, data utility and quality for

analytical goals. Appropriate precautions to prevent re-identification will be taken. Records of the performed privacy protection technique will be maintained for each dataset as will be duly described in Local DMP.

In the project, any access to such data will be granted only to authorised partners for data handling. Furthermore, access for information or data input (even change) will also be restricted only to authorised users to ensure their confidentiality and reserved only for these partners that collect and provide data.

Lawful, ethical, and compliance with FAIR, for all personal data, is under the full responsibility of the partner who owns the data (or if appropriate, the process in which the data is collected). Appropriate procedures and routines shall be included in the Local DMP or policy document.

Should sensitive personal data be collected (or processed) the routines outlined in the deliverables of WP13 Ethics Requirements shall be applied.

Data repositories

Data for scientific use and re-use may be available in open access databases, restricted access databases, and in databases provided by partner or by external actors (Category B data in this document). There are a huge number of such repositories where datasets for AI and machine learning are made available Repositories may be specialised in for example biological or geological fields, or they may be of more general character with regard to scientific fields and more narrow with regard to technology. A number of datasets with social media content, like Twitter and Facebook, Flickr images, blogs, posts, etc. are publicly available.

The use of data items, even though publicly available, is - as a rule - subject to certain license terms, such as Creative Commons (CC) licenses. The reproduction, making available, modification of copyrighted content qualifies as copyright relevant actions and, in principle, require authorization by the right holder, unless exceptions apply.

These basic rights and license terms are attached and communicated via metadata. The licensing and access restrictions are - to reasonable extent - enforced by the technical infrastructure. Each partner is encouraged to include routines for this in their Local DMP.

Data generated for research purposes

Category C generated data that is not personal data is named “other data”. For this data, in general, the Local DMP and the FAIR principles are applicable. Depending on the kind of data and the circumstances, national and/ or EU regulation may be applicable. It is advised that the organisation DPO is consulted regularly and when in doubt.

Allocation of resources

TAILOR funds can be used by the partners to cover costs associated with assuring compliance with the Open Access and FAIR policies (cost category D3 in the Budget).

Data security

Partners will give special attention to the confidentiality of data storage and processing. They will commit to implement all appropriate technical and organisational measures necessary to protect potential personal data against accidental or unlawful destruction or accidental loss, alteration, unauthorised disclosure or access, and against all other unlawful forms of processing, taking into account the particularity of the performed processing operations. Local DMPs are expected to adequately cover data security matters, and organisational routines and technical infrastructure are likewise expected to be appropriate.

DMP relation to ethics deliverables

WP13 Ethics requirements has four deliverables covering the following items, which are therefor not, or only very briefly, covered in this DMP:
D13.1: Procedures should children and/or adults unable to give informed consent be involved.
D13.2 : Informed consent procedures to identify/recruit research participants, and incidental findings policy.
D13.3 : Procedures for legal compliance if personal data is to be transferred between the EU and a non-EU country or international organisation, and procedures if research involves profiling.
D13.4 : Risk assessment and details on measures to prevent misuse of research findings.

References

About	link
Charter of fundamental rights of the EU	https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:12012P/TXT&from=EN
H2020 online manual on data management	https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm
EU Data Protections Rules	https://ec.europa.eu/info/law/law-topic/data-protection/eu-data-protection-rules_en
What is sensitive personal data?	https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/legal-grounds-processing-data/sensitive-data/what-personal-data-considered-sensitive_en
EU AI Alliance, official documents and reports	https://futurium.ec.europa.eu/en/european-ai-alliance/pages/official-documents-and-reports
AI HLEG	https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

ALTAI	https://futurium.ec.europa.eu/en/european-ai-alliance/pages/altai-assessment-list-trustworthy-artificial-intelligence
Guidelines on FAIR Data Management in Horizon 2020	https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf
Nature publication on FAIR principles	https://www.nature.com/articles/sdata201618
EU on Ethics and Data protection	https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/ethics/h2020_hi_ethics-data-protection_en.pdf

Appendix 1: Guidance for Local DMP

A DMP has been produced for TAILOR, and this appendix is provided to support in the development of Local DMPs. There may be specific issues for individual datasets (e.g. regarding openness), and individual DMPs for each data-set may be preferred. Guidelines on FAIR Data Management in Horizon 2020 are available in the H2020 Online Manual.

FAIR data management
In general terms, your research data should be 'FAIR', that is findable, accessible, interoperable and re-usable.

More information about FAIR:
FAIR data principles (FORCE11 discussion forum)
FAIR principles (article in Nature)

This table provides a summary of the Data Management Plan (DMP) issues to be addressed.

DMP component	Issues to be addressed
1. Data summary	· 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
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 mechanisms. 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 type of metadata will be created and how

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
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?
2.4. Increase data re-use (through clarifying licences)	· Specify how the data will be licenced to permit the widest reuse possible · Specify when the data will be made available for reuse. 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 usable 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
3. Allocation of resources	· 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
4. Data security	· Address data recovery as well as secure storage and transfer of sensitive data
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
6. Other	· Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any)

Appendix 2: DMP questionnaire template

Local DMPs are obligatory for all TAILOR partners. All partner PIs will be required to fill in a form based on the questions in the table below. The questionnaire may be modified as the project develops.

My name	
My organisation (ID, acronym)	
Email to organisation Data Protection Officer (DPO)	
Organisation website for data protection matters, if public.	
Questions about Local Data Management Plan related to and relevant for our involvement in TAILOR	
Local Data Management Plan (DMP) established (Y/N) and	
includes routines for Data Processing Impact Assessment (DPIA) (Y/N)	
includes routines for assurance of FAIR(findable, accessible, inter-operable, re-usable) compliance (Y/N)	
if No on the above question, is FAIR policies covered in another official document in your organisation? (Y/N)	
Local DMP, FAIR policies, and other policy documents cover adequately the TAILOR related actions that my organisation is involved with (Y/N)	
Please state name and provide web addresses (if public) to the documents and routines that your refer to for DMP and FAIR compliance	